Enhancing Skin Diseases Classification Through Dual Ensemble Learning and Pre-trained CNNs

Oussama El Gannour[1], Soufiane Hamida[2], Yasser Lamalem[3], Bouchaib Cherradi[4], Shawki Saleh[5], Abdelhadi Raihani[6]
EEIS Laboratory-ENSET of Mohammedia, Hassan II University of Casablanca, Mohammedia, Morocco[1,2,4,5,6]
GENIUS Laboratory, SupMTI of Rabat, Rabat, Morocco[2]
Computer Research Laboratory (L@RI), Ibn Tofail University of Kenitra, Kenitra, Morocco[3]
STIE Team, CRMEF Casablanca-Settat, El Jadida, Morocco[4]

***Abstract*—Skin diseases represent a variety of disorders that can affect the skin. In fact, early diagnosis plays a central role in the treatment of this type of disease. This scholarly article introduces a novel approach to classifying skin diseases by leveraging two ensemble learning techniques, encompassing multi-modal and multi-task methodologies. The proposed classifier integrates diverse information sources, including skin lesion images and patient-specific data, aiming to enhance the accuracy of disease classification. By simultaneously utilizing image input and structured data input, the multi-task functionality of the classifier enables efficient disease classification. The integration of multi-modal and multi-task techniques allows for a comprehensive analysis of skin diseases, leading to improved classification performance and a more holistic understanding of the underlying factors influencing disease diagnosis. The efficacy of the classifier was assessed using the ISIC 2018 dataset, which comprises both image and clinical information for each patient with skin diseases. The dataset used in this study comprises images of seven different types of skin diseases and their associated medical information. The findings of our proposed approach show that it outperforms traditional single-modal and single-task classifiers. The results of this study demonstrate that the proposed model attained an accuracy of 97.66% for the initial classification task (image classification). Additionally, the second classification task (clinical data classification) achieved an accuracy of 94.40%.**

Keywords*—Multi-modal approach; multi-task approach; transfer learning; deep learning; skin diseases classification

I. INTRODUCTION

Skin diseases, also known as dermatological conditions [1], represent a diverse group of disorders that can affect the skin, hair, nails, and mucous membranes. These disorders exhibit a wide spectrum of clinical presentations, ranging from benign and self-limiting conditions [2], such as common warts and seborrheic dermatitis, to severe and debilitating conditions, such as bullous pemphigoid and cutaneous lymphoma. The classification of skin diseases is crucial for proper diagnosis, treatment, and management. Accurate and timely diagnosis allows for the initiation of appropriate therapy and can prevent the progression of the disease. Moreover, the classification of skin diseases is a challenging task [3], that requires a thorough understanding of the clinical presentation, histopathological features, and laboratory results. In addition to this, a misdiagnosis or delayed diagnosis can lead to inappropriate treatment [4], which can result in poor outcomes, increased morbidity, and increased healthcare costs. The use of Artificial Intelligence (AI) techniques for the classification of skin lesions and other diseases, including COVID-19 [5] and diabetes diseases [6], [7], have gained increasing attention in recent years [8], due to their potential to improve diagnostic accuracy and reduce the time required for diagnosis.

AI techniques are based on the ability of algorithms to learn from data and make predictions based on that learning. These techniques can be used to analyze large amounts of data and identify patterns [9], [10], that may be difficult for human experts to discern. In the context of skin disease classification, machine learning and deep learning techniques [11], [12], can be used to analyze images of skin lesions, identify patterns in clinical data, and analyze laboratory results. The utilization of these algorithms in skin disease classification has the potential to improve diagnostic accuracy and reduce the time required for diagnosis. However, the development and implementation of AI techniques for skin disease classification [13], requires a thorough understanding of the data and the algorithms used, as well as a rigorous evaluation of their performance. Recent advancements in AI have also had a significant impact on various other fields [14]–[16].

In the field of skin disease classification, previous studies have primarily focused on using deep learning techniques. While these approaches have shown promising results, there are significant gaps that necessitate further research. One notable gap is the limited exploration of multi-modal and multi-task deep neural networks for enhanced skin disease classification. Existing studies have predominantly focused on single-modal data or single-task models, neglecting the potential benefits of integrating diverse data modalities or leveraging multiple related tasks simultaneously. By addressing this gap, our research aims to develop a transfer learning approach in the context of multi-modal and multi-task deep neural networks, enabling more comprehensive and accurate skin disease classification. This approach holds the potential to leverage the complementary information from various data modalities, such as images, genetic data, and patient history, to improve the overall classification performance. Additionally, the inclusion of multiple related tasks, such as disease severity assessment or treatment recommendation, can further enhance the diagnostic capabilities of the system. By bridging this gap and exploring the potential of multi-modal and multi-task deep neural networks, our research contributes to the advancement of skin disease classification techniques and has practical implications for improving diagnostic accuracy and patient care.

In this study, a multi-modal, multi-task DNN classifier is used to introduce an efficient MDSS [17], [18], that can identify skin conditions in dermatoscopic pictures. The innovative idea is to combine the outputs of two artificial intelligence models based on the TL approach and DNN models [19], [20], using multi-model and multi-task techniques. To build an accurate model compared to the published approaches at the time. To improve the performance of skin disease categorization, this model has been used to integrate medical information, such as past medical history and laboratory findings, with imagery data. This research used the ISIC 2018 dataset to train and evaluate our proposed MDSS, which contains photos of seven distinct types of skin disorders. Furthermore, in addition to the photos provided, the dataset used includes critical clinical information about each image [21], such as age, gender, and location. Hence, the main objective of our article is to enhance the accuracy of skin disease classification, leading to improved diagnostic outcomes, timely interventions, and better patient care. By providing practical insights and demonstrating the applicability of our approach, we aim to empower healthcare practitioners with a reliable tool [22], that can enhance their diagnostic capabilities and positively impact patient outcomes.

This manuscript is structured as follows: A thorough examination of prior research in the domain of skin disease classification utilizing DL techniques is provided in Section II. The design and implementation of the proposed Multi-Modal and Multi-Task classifier, as well as the methodology employed, are expounded upon in Section III. The results of the experiments, including a performance assessment, are detailed in Section IV. Finally, in Section V, the conclusion of this study and potential areas for future research are outlined.

II. Related Works

Deep learning [23], a subset of machine learning [24], has been increasingly used in the field of skin disease classification in recent years. This is due to its ability to automatically extract features [25], and patterns from large and complex datasets, such as images of skin lesions. In this section, we will review the current state-of-the-art in skin disease classification over the last few years using DL techniques. We will examine the various techniques that have been proposed in these studies, such as TL and DNNs, to compare these methods with our proposed classifier in this manuscript.

A study presented in [26], explores the use of DL techniques for the classification of skin lesions on imbalanced small datasets. The authors propose the use of a single model of DL and evaluate its performance in comparison to traditional machine learning methods and human experts. They found that this approach has the potential to improve diagnostic accuracy and reduce the time required for diagnosis, even when working with small, imbalanced datasets. However, the use of DL on imbalanced small datasets also poses challenges such as overfitting and a lack of robustness in the classifier. The authors also suggest potential directions for future research in this field to overcome these challenges, such as the development of more advanced DL architectures and techniques and the integration of additional clinical data. The best proposed model in this study, namely RegNetY-3.2G-Drop, achieved a balanced accuracy value of 85.8% using the ISIC 2018 dataset.

A scientific study featured in [27], presents a new method for skin lesion classification using DL techniques. The proposed method, called SSD-KD, is a self-supervised, diverse knowledge distillation method that uses a lightweight model to classify skin lesions from thermoscopic images. The authors evaluate the performance of this method and show that it can achieve an accuracy of 84.6% and generalization capability even when working with a small dataset. The authors also point out that this approach is an efficient method for skin lesion classification, especially when there is a lack of labeled data. On the other hand, one of the limitations of this study is that it achieves a low level of accuracy.

A scholarly article published in [28], introduces a new technique for diagnosing malignant melanoma using DL techniques. The proposed method, called 2-HDCNN, is a two-tier hybrid dual convolutional neural network feature fusion approach that fuses multiple features from different sources to improve diagnostic accuracy. The authors of the article have evaluated the performance of this method and have found that it is able to achieve high accuracy and generalization capability on the task of malignant melanoma diagnosis, with an accuracy of 92.15%. This means that the method can accurately diagnose malignant melanoma in a high percentage of cases.

A paper appearing in [29], provides a new method for extracting and classifying skin lesion features using DL techniques. The proposed method uses regularization techniques to improve the accuracy and robustness of the model. It also uses layer-wise weight norm-based feature extraction to extract informative features from the skin lesion images. The authors evaluate the performance of this method on several datasets and show that it can achieve an accuracy of 94.42% on the ISIC 2018 dataset, 91.73% accuracy on the ISIC 2019 dataset, and 93% accuracy when evaluated on the combined dataset.

A work documented in [30], discusses a novel approach for skin lesion classification using DL techniques. The proposed method, called End-to-End Decoupled Training (E2EDT), is designed to handle the long-tailed distribution problem, which is a common issue in the skin lesion classification task. E2EDT is a robust DL method that decouples the training process into two stages: pre-training and fine-tuning. The authors evaluate the performance of this method using the ISIC 2018 dataset and show that it can achieve a balanced accuracy of 87%.

Table I presents a comprehensive overview of the reviewed studies, highlighting essential information such as the number of classes, the implemented classifiers, and the highest performance attained based on the evaluation metrics employed. This table acts as a valuable resource, providing a condensed overview of the essential findings from the literature review. By presenting this information in a concise and structured format, Table I facilitates easy reference and comparison of the different approaches and outcomes reported in the studies. This condensed representation of essential findings promotes efficient data analysis and aids in the identification of research gaps and future directions.

TABLE I. SUMMARY OF SOME LITERATURE WORK

Ref	Datasets (Classes)	Models	Accuracy
[26]	ISIC 2017 Dataset (3)	RegNetY-3.2G-Drop	74.6%
	ISIC 2018 Dataset (7)		85.8%
	ISIC 2019 Dataset (8)		59.3%
	7-PT Dataset (5)		65.7%
[27]	ISIC 2019 Dataset (8)	SSD-KD	84.6%
[28]	ISIC-2018 Dataset (2)	2-HDCNN	92.15%
[29]	ISIC 2018 Dataset (2)	CLCM-net	94.42%
	ISIC 2019 Dataset (2)		91.73%
[30]	ISIC 2018 Dataset (7)	E2EDT	87%

In recent years, deep learning techniques have gained significant popularity for skin disease classification. However, this approach is not without limitations. One primary limitation lies in the availability of labeled data, as deep learning models necessitate large amounts of accurately labeled data for effective training. Acquiring such data, especially for rare skin diseases, can be challenging, which hinders the development and deployment of deep learning models in these cases. Another limitation pertains to the quality of the available data. The performance of deep learning models heavily relies on the quality of the input data. Poor data quality, characterized by noise, missing information, or inconsistencies, can negatively impact the model's performance, leading to overfitting and limited generalization capabilities. Furthermore, deep learning models may struggle to generalize well to unseen data, particularly when the distribution of the new data differs significantly from the training data. This limitation poses challenges in real-world scenarios where the model needs to perform accurately in diverse and previously unseen cases. Additionally, the existing literature predominantly focuses on single-modal data or single-task models, neglecting the potential advantages of incorporating multi-modal data and leveraging multi-task learning approaches. By solely considering a single data modality or a specific task, previous studies failed to exploit the complementary information present in diverse data sources, which could potentially enhance the classification performance and broaden the scope of skin disease diagnosis.

Addressing these limitations and exploring the potential of multi-modal data and multi-task learning in the context of skin disease classification is crucial to overcome the data scarcity challenge, improve generalization capabilities, and advance the state-of-the-art in this field.

III. MATERIALS AND METHODS

A. *Schematic of the Planned Multi-Modal and Multi-Task Classifier for Skin Disease Monitoring*

In this subsection, a comprehensive approach for building a multi-modal and multi-task classifier for skin diseases is presented. The proposed framework leverages the EfficientNetV2L network for processing the image dataset and the DNN for handling the clinical dataset. The following paragraphs provide additional details, examples, and techniques to further elucidate the process.

To initiate the classifier's development, a large dataset of skin lesion images and corresponding clinical data is collected. In this study, the ISIC 2018 dataset is specifically chosen due to its widespread utilization and its potential as a valuable resource for constructing the multi-modal and multi-task classifier. The ISIC 2018 dataset comprises a diverse range of skin lesion images along with associated clinical information, enabling the model to learn from both visual and clinical data sources. On the other hand, pre-processing of the collected data is then conducted to ensure its suitability for model training. This crucial step involves data cleaning techniques such as eliminating noise, addressing data inconsistencies, and handling missing or irrelevant information. Furthermore, techniques such as normalization can be employed to enhance the quality of the image dataset, while feature scaling methods can be applied to the clinical dataset. Following data pre-processing, the classifier extracts meaningful features from the image data and the clinical data. The image data is processed using the EfficientNet V2L network, which has demonstrated its effectiveness in image classification tasks. The network employs advanced techniques such as convolutional layers and attention mechanisms [31], [32], to capture intricate visual patterns and extract discriminative features from the skin lesion images. Similarly, the clinical data is fed into a deep neural network (DNN) architecture that is specifically designed to handle heterogeneous clinical data, extracting relevant features from patient-specific attributes, laboratory results, and other clinical information. The extracted features from both modalities are then integrated within a multi-task learning framework. This framework allows the model to simultaneously learn from both image and clinical data, leveraging the complementary information contained in each modality. A shared encoder architecture is employed to concatenate the outputs of the EfficientNet V2L network and the DNN, combining the learned representations from both modalities into a unified feature representation. This fused representation is subsequently passed through a final classifier layer, enabling the model to make accurate predictions regarding skin disease classification. To assess the performance of the proposed multi-modal and multi-task classifier, a separate test dataset is utilized. This dataset comprises skin lesion images and corresponding clinical data that were not involved in the training process. By evaluating the classifier on this independent dataset, its ability to accurately diagnose skin diseases can be quantitatively measured, providing insights into its efficacy and generalization capability.

The flowchart (Fig. 1) accompanying the text visually illustrates the different stages and processes involved in the design and implementation of the proposed multi-modal and multi-task classifier. This flowchart serves as a comprehensive guide, offering a clear and organized overview of the sequential steps encompassing data pre-processing, model construction, and model evaluation.

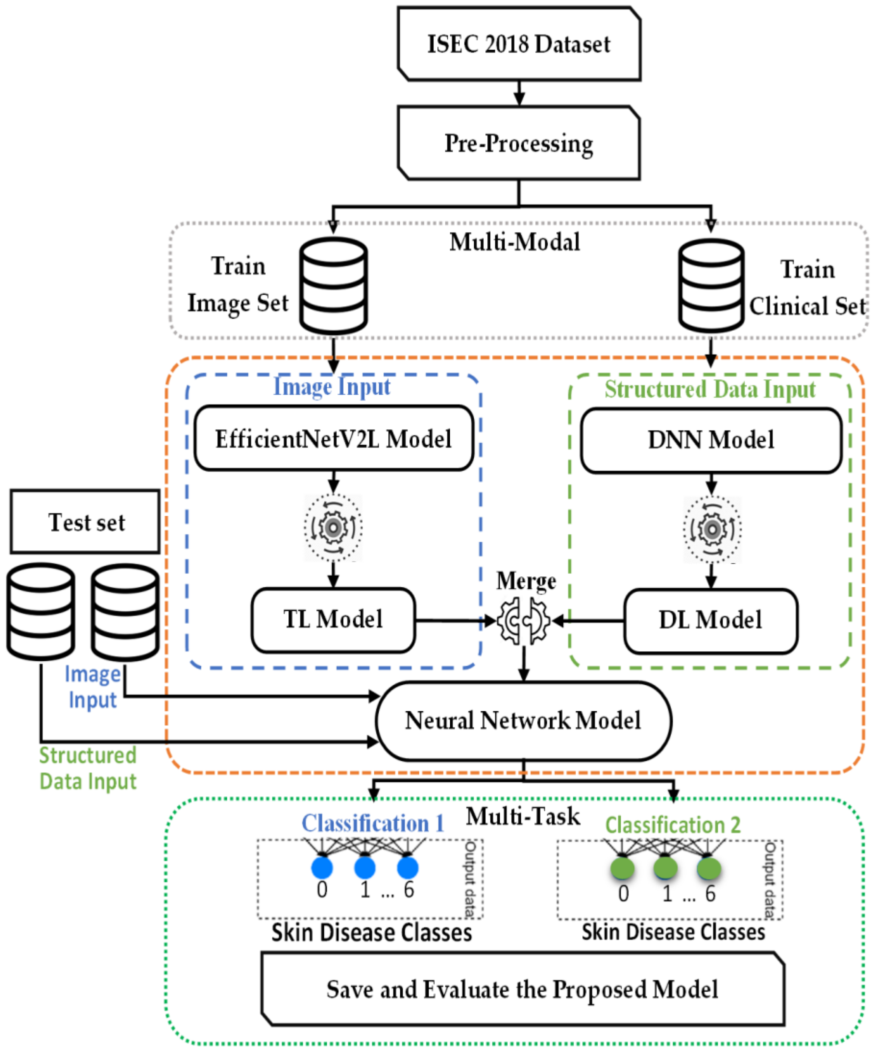
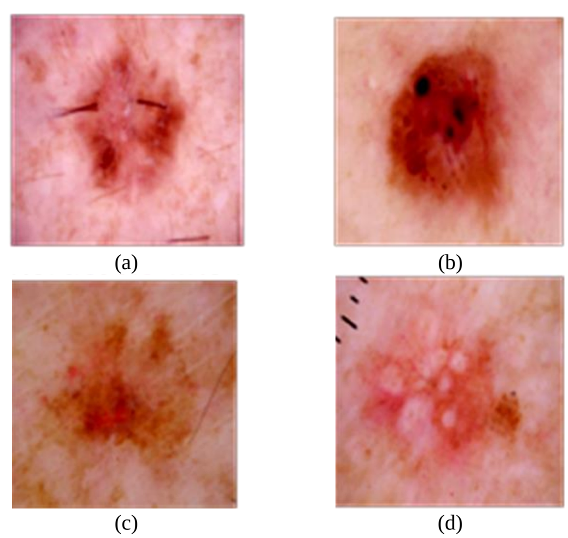
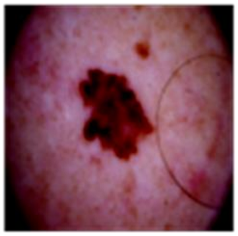


Fig. 1. The architecture of the proposed model.

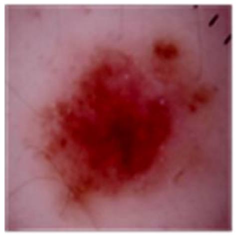
B. ISIC 2018 Dataset Description

The ISIC 2018 dataset is a collection of images of skin lesions, along with corresponding diagnostic information, used for research in the field of skin disease classification using AI techniques. The dataset includes over 10000 images of various skin lesions [33], including malignant melanomas, benign nevi, and other types of skin diseases. The diagnostic information provided with the images includes the diagnosis made by a dermatologist as well as other relevant clinical data. The ISIC 2018 dataset is a widely used benchmark dataset in the field of skin disease classification and is commonly used to evaluate the performance of new AI algorithms and models. As depicted in Fig. 2, several examples are presented from the ISIC 2018 dataset.





(e)



(f)

(g)

Fig. 2. The examples of the seven different classes of skin disease present in the dataset : (a) Actinic keratoses; (b) Basal cell carcinoma; (c) Benign keratosis-like lesions; (d) Dermatofibroma; (e) Melanoma; (f) Melanocytic nevi; and (g) Vascular lesions.

The dataset used in this study comprises images of 7 different types of skin diseases [34], including Actinic keratoses, Basal cell carcinoma, Benign keratosis-like lesions, Dermatofibroma, Melanoma, Melanocytic nevi, and Vascular lesions. Moreover, alongside the provided images, the dataset used encompasses essential clinical information pertaining to each picture, encompassing age, gender, localization, and Dx_type (denoting the modality through which the skin disease was diagnosed). The dataset is divided into a training set and a test set, with 7500 images in the training set in addition to their clinical data and 2500 images in the test set with their associated medical information, respectively.

C. Deep Learning Algorithms for Classification

a) Multi-Modal and Multi-Task Neural Network: A Multi-Modal and Multi-Task Neural Network is a type of DL architecture that can process and analyze multiple types of input data [35], such as images, text, and audio, simultaneously. This type of network can perform multiple tasks, such as classification and segmentation, using the same network architecture, allowing for more efficient and effective processing of data. Additionally, this type of network can improve the performance of each task [36], by leveraging the shared representations and features learned from the other tasks. This is done by sharing the weights of the network across different tasks, which can help improve the generalization of the model and reduce the need for additional training data. Overall, multi-modal, and multi-task networks are well-suited for applications where multiple types of data need to be analyzed together, such as medical imaging and clinical data. The diagram presented in Fig. 3 illustrates a multi-modal and multi-task neural network model, which is represented through its architectural components and connections.

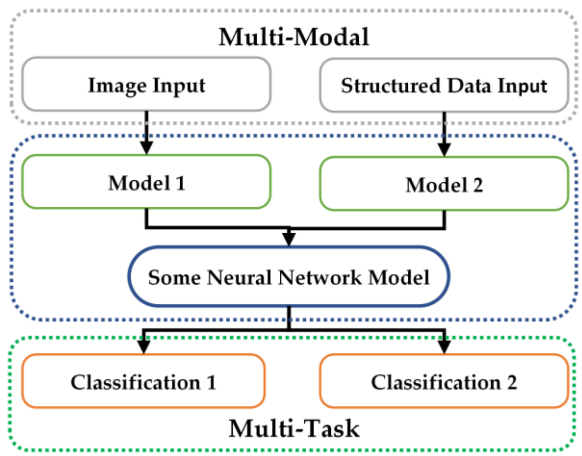


Fig. 3. An example of a multi-modal and multi-task neural network.

b) EfficientNet V2L Network: EfficientNet V2L is a type of neural network architecture [37], that is designed to improve the efficiency and effectiveness of DL models. This network is an updated version of the original EfficientNet, which was designed to improve the performance of DL models while reducing their computational complexity. The V2L version of EfficientNet is specifically designed to handle large-scale image datasets, and it can achieve state-of-the-art performance on a wide range of image classification tasks. The network is characterized by its unique architecture, which includes depth-wise separable convolutions, linear bottlenecks, and efficient scaling of network depth and width. These design choices allow the network to achieve high performance while using fewer parameters than other state-of-the-art architectures. EfficientNetV2 models utilize a scaling method and a neural architecture search algorithm to optimize both the architecture and hyperparameters of the network simultaneously. These models are specifically designed for image classification tasks and come in a range of sizes, referred to as "scales." Each scale corresponds to a distinct model architecture and number of parameters, with larger scales possessing a greater number of parameters and higher accuracy. The architecture search space has been expanded to include novel operations, such as FusedMBConv, in addition to conventional CNN operations.

c) Deep Neural Network: DNNs are a class of machine learning models [38], that are composed of multiple layers of artificial neurons. These layers are interconnected and process the input data through a series of mathematical computations known as activation functions. DNNs are trained to learn patterns and features in the input data, allowing them to perform tasks such as image recognition, natural language processing, and speech recognition. DNNs are known for their ability to represent complex, non-linear relationships between input and output and can often achieve state-of-the-art performance on a wide range of tasks. They are widely used in many medical fields and have become a popular tool for solving complex problems.

D. Confusion Matrix and Measures

A confusion matrix is a table that is often used to describe the performance of a classification algorithm, specifically in the context of medical diagnosis. In the case of skin diseases, a confusion matrix [39], would display the number of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) for a given diagnostic test. These values can then be used to calculate various performance measures [40], such as accuracy, precision, recall, and specificity, which can give insight into the effectiveness of the diagnostic test. The following formulas compute these metrics:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{FN+TP}} \quad (3)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \quad (4)$$

IV. Findings and Discussion

In the current section, the primary experimental results of the scientific study are presented. Initially, the training performance of all utilized models is depicted through plots that depict the evolution of accuracy and loss over time. Furthermore, a confusion matrix is presented to provide a thorough understanding of the classification outcomes.

A. Implementation Platform

The present study carried out all experiments using the Python programming language and the Jupyter Lab environment. The AI models were trained on the Google Colab platform, which is a Jupyter Notebook-based cloud service for researching and training AI algorithms. The computational resources employed in this setup included a Tesla K80 GPU with 12GB of GDDR5 VRAM, an Intel Xeon Processor with two cores operating at 2.20 GHz, and 13 GB of RAM. The construction of models was facilitated through the utilization of two primary implementation libraries: Keras and Autokeras. Additionally, the Adam optimizer and cross-entropy loss function [41], were employed for training all algorithms. Table II presents the hyperparameters used for fine-tuning the pre-trained models utilized in this research.

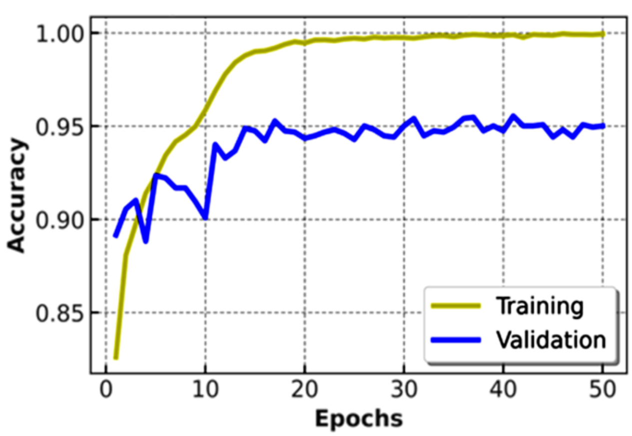
TABLE II. The Parameters and Functions Utilized in the Training Procedures

Network	Epochs	Batch Size	Loss Function	Optimizer	Learning Rate
Proposed Model	50	16	Categorical Cross entropy	Adam	1e-4

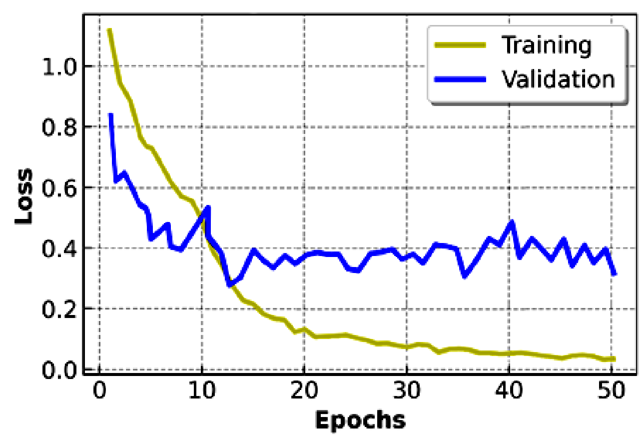
B. Experimental Results

a) Training of DL Models Results: The training results of the proposed multi-modal and multi-task DNN with TL were evaluated using accuracy, precision, recall, and specificity metrics. These results were also evaluated using learning curves, which showed a consistent improvement in performance as the number of training epochs increased. The learning curves also indicate that the proposed approach can effectively utilize the additional information provided by this proposed approach. The loss curve, on the other hand, shows a decrease, indicating that the model is learning and generalizing well. The convergence of both curves suggests that the proposed approach effectively leverages multi-modality and multi-task information, resulting in a robust classifier for skin disease classification. Generally, these results demonstrate the effectiveness of the proposed approach in improving the classification of skin diseases. Fig. 4 illustrates the accuracy and loss curves of the multi-modal and multi-task DNN with a TL classifier during the training phase.

b) Testing of DL Models Results: The effectiveness of the models was evaluated in this study using a separate set of test data that incorporated both image and clinical data inputs. The performance of the model for each class of skin disease was visualized through the generation of a confusion matrix, which subsequently informed the calculation of performance metrics. The results of the multi-class classification confusion matrix are presented in Fig. 5.

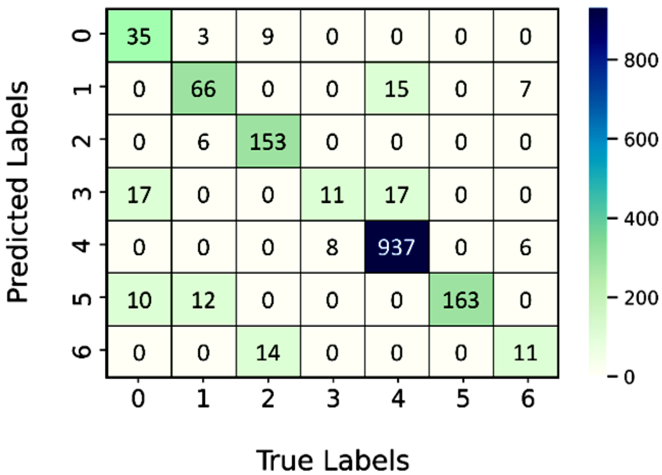


(a)

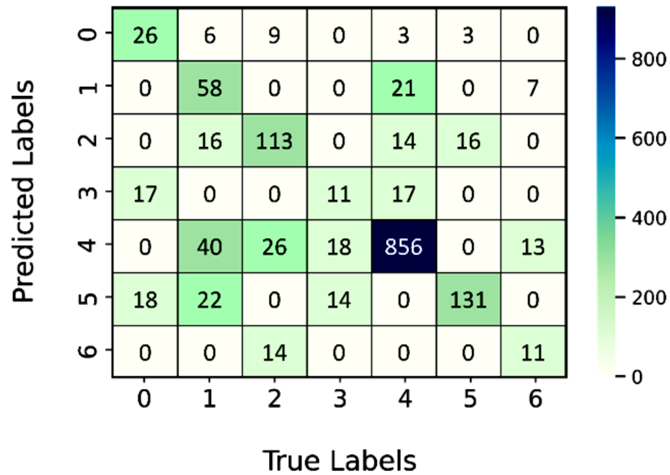


(b)

Fig. 4. Performance of the multi-modal and multi-task DNN with the TL classifier : (a) The accuracy of the model over time; (b) The decrease in loss function over time during the training phase.



(a)



(b)

Fig. 5. Confusion matrix illustrates the performance of the proposed classifier: (a) Initial classification task (image input); (b) Second classification task (clinical data input).

These matrices show that the model performed well for most of the classes, with a high number of TP and a low number of FP and FN. The confusion matrix also revealed some misclassification for a few classes, which can be further investigated to improve the performance of the model. Overall, these results demonstrate the effectiveness of the proposed approach in classifying skin diseases. Based on the results of the matrices, it can be inferred that the proposed approach for classifying skin diseases is effective. The assessment measures for the performance of the proposed model are also outlined in Tables III and IV.

TABLE III. PERFORMANCE ANALYSIS OF THE INITIAL CLASSIFICATION TASK USING METRICS: ACCURACY, PRECISION, RECALL, AND SPECIFICITY

Class	**Accuracy**	**Precision**	**Recall**	**Specificity**
0	97.43%	56.45%	74.47%	98.16%
1	97.16%	75.86%	75.00%	98.53%
2	98.09%	86.93%	96.23%	98.31%
3	97.23%	57.89%	24.44%	99.46%
4	96.97%	96.75%	98.55%	94.17%
5	98.55%	100.00%	88.11%	100.00%
6	98.22%	45.83%	44.00%	99.13%
Average	**97.66%**	**74.25%**	**71.54%**	**98.25%**

TABLE IV. PERFORMANCE ANALYSIS OF THE SECOND CLASSIFICATION TASK USING METRICS: ACCURACY, PRECISION, RECALL, AND SPECIFICITY

Class	**Accuracy**	**Precision**	**Recall**	**Specificity**
0	96.27%	42.62%	55.32%	97.59%
1	92.53%	40.85%	67.44%	94.06%
2	93.67%	69.75%	71.07%	96.35%
3	95.60%	25.58%	24.44%	97.80%
4	89.87%	93.96%	89.82%	89.95%
5	95.13%	87.33%	70.81%	98.56%
6	97.73%	35.48%	44.00%	98.64%
Average	**94.40%**	**56.51%**	**60.42%**	**96.13%**

The tables provide numerical values for each performance metric, which allows for a quantitative assessment of the multi-task classification performances. The numerical values represent the specific performance of the model for each metric, and by comparing these values to other metrics, we can understand how well the model is performing.

C. Discussion

A novel methodology for skin disease classification is presented that incorporates the utilization of multi-modal and multi-task classifier. The proposed classifier integrates multiple modalities of data, including visual representations of skin lesions and patient-specific characteristics, to enhance the precision of disease classification. The multi-task aspect of the classifier enables it to concurrently classify the disease through the integration of image-based inputs and structured data. The proposed approach was evaluated using the ISIC 2018 dataset, which includes both image and clinical information for patients with skin diseases. This dataset encompasses data on 7 different categories of skin diseases. Our findings indicate that the proposed model exhibited a high level of accuracy, specifically 97.66% for the primary classification task (Image classification). Furthermore, the second classification task (clinical data classification) demonstrated an accuracy of 94.40%. The results of our proposed methodology demonstrate that it surpasses traditional single-modal and single-task classifiers.

The proposed classifier for improved classification of skin diseases offers several benefits. One of the key contributions of this model is that it utilizes multiple sources of information, including images of skin lesions and patient-specific information, to improve the accuracy of disease classification. Additionally, the multi-task aspect of the classifier allows it to simultaneously classify the disease and predict its severity, providing more comprehensive information for diagnosis. The use of transfer learning techniques also allows for better performance in real-world scenarios and faster training times. While the proposed approach for improved classification of skin diseases offers many benefits, it also has some limitations. One limitation is that the proposed approach is based on a dataset of skin diseases, which means that it may not be able to generalize well to other types of skin diseases or other medical conditions. Another limitation is that the model still requires a large amount of labeled data for training, which can be challenging to obtain. Moreover, while the proposed classifier shows promising findings, it is important to be aware of these limitations and to continue researching ways to improve the model's performance and applicability. Table V presents a comparison of the proposed multi-modal and multi-task DNN with a TL classifier to other methods discussed in the literature. The table compares the performance of our system to other efforts in terms of accuracy metrics. The results show that our proposed system outperforms the other methods discussed in the literature, indicating that it is a promising approach and offers significant potential for improving skin disease classification. The superior performance exhibited by our system reinforces its viability as an advanced and effective method in the field, highlighting its potential for enhancing diagnostic accuracy and facilitating improved patient care.

TABLE V. COMPARATIVE OUTCOMES OF THE PROPOSED APPROACH WITH EARLIER STUDIES PUBLISHED IN THE LITERATURE

Ref	Dataset (Classes)	Models	Accuracy
[26]	ISIC 2018 Dataset (7)	RegNetY-3.2G-Drop	85.8%
[27]	ISIC 2019 Dataset (8)	SSD-KD	84.6%
[30]	ISIC 2018 Dataset (7)	E2EDT	87%
Pr. Model	**ISIC 2018 Dataset (7)**	**M3T_DNN_TL [a]**	**97.66%**
			94.40%

[a] Multi-Modal and Multi-Task DNN with TL

V. CONCLUSION AND PERSPECTIVES

In conclusion, this study proposed a multi-modal and multi-task DNN with TL for improved classification of skin diseases. The proposed approach utilizes multiple sources of information, including images of the skin lesions and patient-specific information, to improve the accuracy of disease classification. The classifier's ability to perform multiple tasks enables it to classify the disease by utilizing both image input and structured data input simultaneously. The results of the study demonstrate that the proposed model achieved high accuracy on the ISIC 2018 dataset, outperforming traditional single-modal and single-task classifiers.

As perspectives, there are several promising avenues for future work building upon the findings of this research. One prospective direction involves exploring additional data sources to further enhance the accuracy and robustness of the model. On the other hand, incorporating genetic data, patient history, or environmental factors could provide a more comprehensive understanding of skin diseases and enable more precise classifications. Additionally, investigating advanced techniques, such as ensemble models and the XAI technique, may improve the performance and generalization capabilities of the classification system.

REFERENCES

[1] P. K. Upadhyay and S. Chandra, “An improved bag of dense features for skin lesion recognition,” Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 3, pp. 520–525, Mar. 2022, doi: 10.1016/j.jksuci.2019.02.007.

[2] G. Yosipovitch, M. Saint Aroman, C. Taieb, F. Sampogna, F. Carballido, and A. Reich, “Skin pain: A symptom to be investigated,” Journal of the American Academy of Dermatology, vol. 88, no. 2, pp. 479–481, Feb. 2023, doi: 10.1016/j.jaad.2022.06.027.

[3] P. Matteucci, R. Pinder, A. Magdum, and P. Stanley, “Accuracy in skin lesion diagnosis and the exclusion of malignancy,” Journal of Plastic, Reconstructive & Aesthetic Surgery, vol. 64, no. 11, pp. 1460–1465, Nov. 2011, doi: 10.1016/j.bjps.2011.06.017.

[4] J. M. Grant-Kels, E. T. Bason, and C. M. Grin, “The misdiagnosis of malignant melanoma,” Journal of the American Academy of Dermatology, vol. 40, no. 4, pp. 539–548, Apr. 1999, doi: 10.1016/S0190-9622(99)70435-4.

[5] S. Hamida, O. El Gannour, B. Cherradi, A. Raihani, H. Moujahid, and H. Ouajji, “A Novel COVID-19 Diagnosis Support System Using the Stacking Approach and Transfer Learning Technique on Chest X-Ray Images,” Journal of Healthcare Engineering, vol. 2021, pp. 1–17, Nov. 2021, doi: 10.1155/2021/9437538.

[6] O. Daanouni, B. Cherradi, and A. Tmiri, “Diabetes Diseases Prediction Using Supervised Machine Learning and Neighbourhood Components Analysis,” in Proceedings of the 3rd International Conference on Networking, Information Systems & Security, Marrakech Morocco: ACM, Mar. 2020, pp. 1–5. doi: 10.1145/3386723.3387887.

[7] O. Daanouni, B. Cherradi, and A. Tmiri, “Self-Attention Mechanism for Diabetic Retinopathy Detection,” in Emerging Trends in ICT for Sustainable Development, M. Ben Ahmed, S. Mellouli, L. Braganca, B. Anouar Abdelhakim, and K. A. Bernadetta, Eds., in Advances in Science, Technology & Innovation. Cham: Springer International Publishing, 2021, pp. 79–88. doi: 10.1007/978-3-030-53440-0_10.

[8] O. Reiter, V. Rotemberg, K. Kose, and A. C. Halpern, “Artificial Intelligence in Skin Cancer,” Curr Derm Rep, vol. 8, no. 3, pp. 133–140, Sep. 2019, doi: 10.1007/s13671-019-00267-0.

[9] S. Khalid, T. Khalil, and S. Nasreen, “A survey of feature selection and feature extraction techniques in machine learning,” in 2014 Science and Information Conference, London, UK: IEEE, Aug. 2014, pp. 372–378. doi: 10.1109/SAI.2014.6918213.

[10] M. A. Mahjoubi, S. Hamida, O. E. Gannour, B. Cherradi, A. E. Abbassi, and A. Raihani, “Improved Multiclass Brain Tumor Detection using Convolutional Neural Networks and Magnetic Resonance Imaging,” IJACSA, vol. 14, no. 3, 2023, doi: 10.14569/IJACSA.2023.0140346.

[11] Ł. Piatek and T. Mroczek, “Analysis and classification of melanocytic skin lesion images,” Procedia Computer Science, vol. 207, pp. 1911–1918, 2022, doi: 10.1016/j.procs.2022.09.249.

[12] B. Zhang et al., “Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning,” Chin. J. Mech. Eng., vol. 34, no. 1, p. 112, Dec. 2021, doi: 10.1186/s10033-021-00629-5.

[13] A. Afroz, R. Zia, A. O. Garcia, M. U. Khan, U. Jilani, and K. M. Ahmed, “Skin lesion classification using machine learning approach: A survey,” in 2022 Global Conference on Wireless and Optical Technologies (GCWOT), Malaga, Spain: IEEE, Feb. 2022, pp. 1–8. doi: 10.1109/GCWOT53057.2022.9772915.

[14] S. Hamida, O. El Gannour, B. Cherradi, H. Ouajji, and A. Raihani, “Handwritten computer science words vocabulary recognition using concatenated convolutional neural networks,” Multimed Tools Appl, Nov. 2022, doi: 10.1007/s11042-022-14105-2.

[15] M.-A. Ouassil, B. Cherradi, S. Hamida, M. Errami, O. E. Gannour, and A. Raihani, “A Fake News Detection System based on Combination of Word Embedded Techniques and Hybrid Deep Learning Model,” IJACSA, vol. 13, no. 10, 2022, doi: 10.14569/IJACSA.2022.0131061.

[16] Y. Lamalem, S. Hamida, Y. Tazouti, O. E. Gannour, K. Housni, and B. Cherradi, “Evaluating multi-state systems reliability with a new improved method,” Bulletin EEI, vol. 11, no. 3, Art. no. 3, Jun. 2022, doi: 10.11591/eei.v11i3.3509.

[17] H. Li, Y. Pan, J. Zhao, and L. Zhang, “Skin disease diagnosis with deep learning: A review,” Neurocomputing, vol. 464, pp. 364–393, Nov. 2021, doi: 10.1016/j.neucom.2021.08.096.

[18] A. Adegun and S. Viriri, “Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art,” Artif Intell Rev, vol. 54, no. 2, pp. 811–841, Feb. 2021, doi: 10.1007/s10462-020-09865-y.

[19] H. K. Kondaveeti and P. Edupuganti, “Skin Cancer Classification using Transfer Learning,” in 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), Buldhana, India: IEEE, Dec. 2020, pp. 1–4. doi: 10.1109/ICATMRI51801.2020.9398388.

[20] M. Lucius et al., “Deep Neural Frameworks Improve the Accuracy of General Practitioners in the Classification of Pigmented Skin Lesions,” Diagnostics, vol. 10, no. 11, p. 969, Nov. 2020, doi: 10.3390/diagnostics10110969.

[21] P. Tang, X. Yan, Y. Nan, S. Xiang, S. Krammer, and T. Lasser, “FusionM4Net: A multi-stage multi-modal learning algorithm for multi-label skin lesion classification,” Medical Image Analysis, vol. 76, p. 102307, Feb. 2022, doi: 10.1016/j.media.2021.102307.

[22] S. Saleh, B. Cherradi, O. El Gannour, N. Gouiza, and O. Bouattane, “Healthcare monitoring system for automatic database management using mobile application in IoT environment,” Bulletin EEI, vol. 12, no. 2, pp. 1055–1068, Apr. 2023, doi: 10.11591/eei.v12i2.4282.

[23] O. El Gannour, B. Cherradi, S. Hamida, M. Jebbari, and A. Raihani, “Screening Medical Face Mask for Coronavirus Prevention using Deep Learning and AutoML,” in 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco: IEEE, Mar. 2022, pp. 1–7. doi: 10.1109/IRASET52964.2022.9737903.

[24] R. Rachidi, M. A. Ouassil, M. Errami, B. Cherradi, S. Hamida, and H. Silkan, “Classifying toxicity in the Arabic Moroccan dialect on Instagram: a machine and deep learning approach,” IJEECS, vol. 31, no. 1, p. 588, Jul. 2023, doi: 10.11591/ijeecs.v31.i1.pp588-598.

[25] S. Hamida, B. Cherradi, O. El Gannour, A. Raihani, and H. Ouajji, “Cursive Arabic handwritten word recognition system using majority voting and k-NN for feature descriptor selection,” Multimed Tools Appl, Mar. 2023, doi: 10.1007/s11042-023-15167-6.

[26] P. Yao et al., “Single Model Deep Learning on Imbalanced Small Datasets for Skin Lesion Classification,” IEEE Trans. Med. Imaging, vol. 41, no. 5, pp. 1242–1254, May 2022, doi: 10.1109/TMI.2021.3136682.

[27] Y. Wang, Y. Wang, J. Cai, T. K. Lee, C. Miao, and Z. J. Wang, “SSD-KD: A self-supervised diverse knowledge distillation method for lightweight skin lesion classification using dermoscopic images,” Medical Image Analysis, vol. 84, p. 102693, Feb. 2023, doi: 10.1016/j.media.2022.102693.

[28] Y. Nancy Jane, S. K. Charanya, M. Amsaprabhaa, P. Jayashanker, and K. Nehemiah H., “2-HDCNN: A two-tier hybrid dual convolution neural network feature fusion approach for diagnosing malignant melanoma,” Computers in Biology and Medicine, vol. 152, p. 106333, Jan. 2023, doi: 10.1016/j.compbiomed.2022.106333.

[29] S. Gopikha and M. Balamurugan, “Regularised Layerwise Weight Norm Based Skin Lesion Features Extraction and Classification,” Computer Systems Science and Engineering, vol. 44, no. 3, pp. 2727–2742, 2023, doi: 10.32604/csse.2023.028609.

[30] A. C. Foahom Gouabou, R. Iguernaissi, J.-L. Damoiseaux, A. Moudafi, and D. Merad, “End-to-End Decoupled Training: A Robust Deep Learning Method for Long-Tailed Classification of Dermoscopic Images for Skin Lesion Classification,” Electronics, vol. 11, no. 20, p. 3275, Oct. 2022, doi: 10.3390/electronics11203275.

[31] O. El Gannour et al., “Concatenation of Pre-Trained Convolutional Neural Networks for Enhanced COVID-19 Screening Using Transfer Learning Technique,” Electronics, vol. 11, no. 1, Art. no. 1, Dec. 2021, doi: 10.3390/electronics11010103.

[32] O. Daanouni, B. Cherradi, and A. Tmiri, “Automatic Detection of Diabetic Retinopathy Using Custom CNN and Grad-CAM,” in Advances on Smart and Soft Computing, F. Saeed, T. Al-Hadhrami, F. Mohammed, and E. Mohammed, Eds., in Advances in Intelligent Systems and Computing, vol. 1188. Singapore: Springer Singapore, 2021, pp. 15–26. doi: 10.1007/978-981-15-6048-4_2.

[33] N. C. F. Codella et al., “Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC),” in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC: IEEE, Apr. 2018, pp. 168–172. doi: 10.1109/ISBI.2018.8363547.

[34] T. Skuhala, V. Trkulja, M. Rimac, A. Dragobratović, and B. Desnica, “Analysis of Types of Skin Lesions and Diseases in Everyday Infectious Disease Practice—How Experienced Are We?,” Life, vol. 12, no. 7, p. 978, Jun. 2022, doi: 10.3390/life12070978.

[35] D. Srihari and P. V., “Multi Modal RGB D Action Recognition with CNN LSTM Ensemble Deep Network,” IJACSA, vol. 11, no. 12, 2020, doi: 10.14569/IJACSA.2020.0111284.

[36] G. Shakah, “Multi-Task Reinforcement Meta-Learning in Neural Networks,” IJACSA, vol. 13, no. 7, 2022, doi: 10.14569/IJACSA.2022.0130734.

[37] V.-T. Hoang and K.-H. Jo, “Practical Analysis on Architecture of EfficientNet,” in 2021 14th International Conference on Human System Interaction (HSI), Gdańsk, Poland: IEEE, Jul. 2021, pp. 1–4. doi: 10.1109/HSI52170.2021.9538782.

[38] M. N. Al-Mhiqani, R. Ahmed, Z. Zainal, and S. N. Isnin, “An Integrated Imbalanced Learning and Deep Neural Network Model for Insider Threat Detection,” IJACSA, vol. 12, no. 1, 2021, doi: 10.14569/IJACSA.2021.0120166.

[39] I. Guyon and A. Elisseeff, “An Introduction to Feature Extraction,” in Feature Extraction, I. Guyon, M. Nikravesh, S. Gunn, and L. A. Zadeh, Eds., in Studies in Fuzziness and Soft Computing, vol. 207. Berlin,

Heidelberg: Springer Berlin Heidelberg, 2006, pp. 1–25. doi: 10.1007/978-3-540-35488-8_1.

[40] H. Dalianis, “Evaluation Metrics and Evaluation,” in Clinical Text Mining, Cham: Springer International Publishing, 2018, pp. 45–53. doi: 10.1007/978-3-319-78503-5_6.

[41] V. S. Suryaa, A. X. A. R, and A. M. S, “Efficient DNN Ensemble for Pneumonia Detection in Chest X-ray Images,” IJACSA, vol. 12, no. 10, 2021, doi: 10.14569/IJACSA.2021.0121084.