ISBN: 978-628-95207-8-1. ISSN: 2414-6390. Digital Object Identifier: 10.18687/LACCEI2024.1.1.940

Web-Based System for the Diagnosis of Skin Lesions Using Deep Convolutional Neural Networks and Transfer Learning Techniques

Elmer Jaén, Bachelor's Degree[1], Zbigniew Omiotek, Doctoral Degree[2], Cristian Pinzón, Doctoral Degree[3]
[1]Technological University of Panama, Panama, elmer.jaen@utp.ac.pa
[2]Lublin University of Technology, Poland, z.omiotek@pollub.pl
[3]Technological University of Panama, Panama, cristian.pinzon@utp.ac.pa

Abstract- ***The diagnosis of skin lesions plays a crucial role in the early detection and treatment of various dermatological conditions. In this study, we present a web-based system for skin lesions diagnosis that utilizes deep learning models to support the identification of six different types of skin lesions (nevus, pigmented benign keratosis, seborrheic keratosis, melanoma, basal cell carcinoma and squamous cell carcinoma). The web application allows users to upload images, which are then processed by the classifier to determine the most likely skin lesion present. Six pre-trained DCNN architectures (VGG16, VGG19, DenseNet201, InceptionV3, MobileNetV2, and Xception) were used in this research. A dataset containing 2400 images was used to train the models. Data augmentation techniques were employed to increase the number of training samples.***

After conducting experimentation and a comprehensive evaluation, we concluded that the deep learning models provided satisfactory results in detecting the different skin lesions. Notably, the VGG16 model exhibited superior classification accuracy (86%) and fast response times, making it the most effective model among the six. The web-based system, designed with a user-friendly and easy-to-use interface serves two purposes: empowering patients to perform self-diagnosis and providing dermatologists with support for more accurate diagnoses. Our findings highlight the potential of deep learning models, particularly the VGG16 architecture, in assisting with the diagnosis of skin lesions. Our work proved that it is possible to build an efficient skin lesions diagnosis tool based on existing web technologies and machine learning methods.

Keywords: Skin lesions classification, deep learning, convolutional neural networks, transfer learning, web-based system.

Corresponding author: *cristian.pinzon@utp.ac.pa

I. INTRODUCTION

Skin lesions are abnormal changes in the appearance or texture of the skin that can be caused by various factors, such as infections, injuries, allergies, genetic disorders, or exposure to environmental agents. Skin lesions can range from benign (harmless) to malignant (cancerous), and can affect different layers of the skin, such as the epidermis (the outermost layer), the dermis (the middle layer), or the subcutis (the innermost layer).

Skin lesions can have significant impacts on the physical and psychological well-being of individuals, as they may cause pain, itching, bleeding, scarring or reduced self-esteem. Moreover, some skin lesions, especially skin cancers, can pose serious threats to life if not diagnosed and treated early. According to the Skin Cancer Foundation, skin cancer is the most common cancer in the United States and worldwide, affecting more than 9,500 people every day and causing more than two deaths every hour in the U.S. alone [1]. The most common types of skin cancer are basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and melanoma. BCC and SCC are collectively known as nonmelanoma skin cancers (NMSCs), which are usually less aggressive and more treatable than melanoma. Melanoma is the most dangerous form of skin cancer, as it can spread rapidly to other organs and cause death and the five-year survival rate for this skin disease is 99% if detected early but drops to 27.3% if detected at a late stage [2].

The diagnosis of skin lesions is usually performed by dermatologists or other trained health professionals using various methods, such as visual inspection, dermoscopy (a technique that uses a magnifying device to examine the skin surface), biopsy (a procedure that involves taking a small sample of the skin tissue for microscopic analysis) [3], or histopathology (a technique that uses a microscope to examine the structure and function of the skin cells) [4].

However, there are some limitations and challenges associated with the conventional methods of skin lesion diagnosis. First, visual inspection and dermoscopy rely heavily on the subjective judgment and experience of the dermatologist, which may vary depending on their level of training and expertise. Second, biopsy and histopathology are invasive procedures that may cause pain, bleeding, infection, or scarring to the patient. Third, there may be a delay in accessing dermatological services due to long waiting times or geographical distances. Fourth, there may be a lack of awareness or education among patients about the importance of seeking medical attention for suspicious skin lesions.

Therefore, there is a need for alternative or complementary methods of skin lesion diagnosis that can overcome some of

these limitations and challenges. One promising approach is to use deep learning models to support the identification of different types of skin lesions from images. Deep learning is a branch of machine learning that uses artificial neural networks to learn from large amounts of data and perform complex tasks. Deep convolutional neural networks (DCNNs) are a type of deep learning model that can extract features from images and perform image recognition and classification [5].

A. *Contribution*

The main contribution of this study is described as follows:

- We present a web-based system with user friendly interface for the diagnosis of skin lesions that uses the VGG16 network architecture.
- The system can identify six different types of skin lesions from images uploaded by users.
- We compared six different pre-trained DCNN architectures and identified the most effective model (VGG16).
- We used a dataset of 2400 images and applied data augmentation techniques to train the models.
- We evaluate the performance and viability of using a web-based system powered by deep learning models for the diagnosis of skin lesions and its potential benefits for both patients and dermatologists.
- For patients, our system could provide a convenient and accessible way to perform self-diagnosis and monitor their skin health.
- For dermatologists, it could provide a reliable and efficient tool to assist them in making more accurate diagnoses.

The rest of this paper is organized as follows: Section 2 reviews the related work on using deep learning for skin lesions classification. Section 3 describes the materials and methods used in this study, including the dataset, the models, and the evaluation metrics. Section 4 presents the system architecture and the technologies used to build the solution. Section 5 shows the results. Section 6 concludes the paper and suggests future work.

II. RELATED WORK

In reference [6], the authors developed a skin lesion classification model using machine learning and convolutional neural network (CNN) techniques. The proposed model was designed to address the issue of highly unbalanced datasets, which is a common problem in skin lesion classification. The HAM10000 dataset was used to train and test the model. The methodology included resizing the images to reduce memory consumption and improve latency, data augmentation to overcome the limited number of images and reduce overfitting, and global feature descriptors to efficiently extract skin lesion features. The proposed CNN model achieved an accuracy of 95.18% in skin lesion classification.

Another common method for skin lesion classification is to use transfer learning, which is a technique that leverages pre-trained models on large datasets and fine-tunes them for specific tasks or domains. Reference [7] proposed an enhanced technique for skin cancer classification using deep convolutional neural network (CNN) with transfer learning models. The authors applied pre-processing steps such as resizing, normalization, and data augmentation to improve the accuracy of the model. They used transfer learning to fine-tune pre-trained CNN models and achieved high accuracy in binary classification of benign and malignant skin lesions. Their model achieved a testing accuracy of 94.27%. However, it is noted that while their model performed well in binary classification, it may not be the most suitable approach when the specific type of skin lesion needs to be identified for appropriate treatment and management.

Reference [8] developed an all-inclusive smartphone application for skin cancer diagnostics. The application captures or imports images of skin lesions, and classifies malignancy based on a support vector machine (SVM) classifier. The approach is made computationally light and user-friendly by using adaptive algorithms in the individual data-processing stages. However, their model was trained on a relatively small dataset (200 images), which may limit its performance compared to models trained on larger datasets. Additionally, their system was designed to classify skin lesions as either suspicious of melanoma or benign, therefore it will not be appropriate to detect other types of skin lesions.

In addition, ref. [9] proposed a method for classifying skin lesions using transfer learning and pre-trained deep neural network GoogleNet. The proposed method involves replacing several layers of the GoogleNet architecture to adapt it to classify skin images. The model was trained on the ISIC 2019 dataset. The proposed method achieved an accuracy of 94.92%.

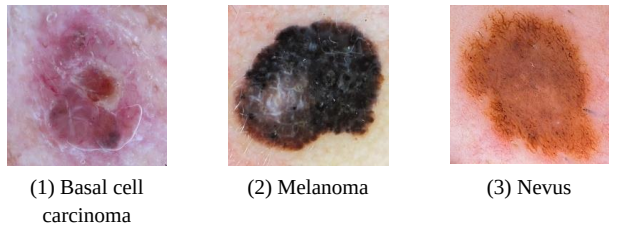
In summary, previous studies have used various methods for skin lesion classification, such as CNNs, transfer learning, and SVMs. However, most of these methods have some limitations, such as the use of small or unbalanced datasets, the focus on binary classification rather than multi-class classification, or the disregard of user experience and accessibility. In this study, we aim to overcome some of these limitations by developing and evaluating a web-based user-friendly and easy-to-use diagnosis tool that employs the VGG16 network architecture to support the detection of six different types of skin lesions.

III. MATERIALS AND METHODS

This project falls within the realm of applied research, focusing on providing a descriptive analysis of a real problem and proposing a potential solution. The following section presents the step-by-step process followed in executing this project:

A. Dataset

The data for this study was obtained from the ISIC dataset [10], which contains images of various types of skin lesions. To focus on the region of interest and to remove any irrelevant information, the images were cropped using a segmentation algorithm. A total of 2400 images were used to train the models, with 400 images per class for six classes of skin lesions. The dataset was divided into three subsets: 50% for training (200 images per class), 25% for testing (100 images per class), and 25% for validation (100 images per class).



(1) Basal cell carcinoma (2) Melanoma (3) Nevus

(4) Pigmented benign keratosis (5) Squamous cell carcinoma (6) Seborrheic keratosis

Figure 1. Sample images for the six skin lesions selected in this study.

B. Classification Models

In this study, six pre-trained DCNN architectures were used for skin lesion classification: VGG16 [11], VGG19 [11], DenseNet201 [12], InceptionV3 [13], MobileNetV2 [14], and Xception [15]. These models have been trained on the ImageNet dataset [16] and have achieved state-of-the-art results on various image recognition tasks.

As part of our experiment, the models have been built as sequential Keras models, which are composed of a linear stack of layers. The models consisted of a convolutional base, responsible for feature extraction from the input images, followed by our custom classifier, which makes predictions based on the extracted features. The first layer of our classifier was a Flatten layer, which reduces the multi-dimensional output of the convolutional base into a one-dimensional vector to be used by the subsequent layers of the classifier.

Subsequently, two densely connected layers were incorporated in the classifier. Each neuron in these layers was fully connected to neurons in the preceding and succeeding layers. The first dense layer consisted of 256 neurons and utilized the ReLU (Rectified Linear Unit) activation function. ReLU returns the input value if positive and zero otherwise, facilitating nonlinear transformations in the network [17].

The second dense layer served as the output layer, consisting of six neurons corresponding to the possible classes for prediction. This layer employed the softmax activation function, enabling the conversion of the neuron outputs into a probability distribution across the classes.

One dropout layer with rate = 0.5, which randomly sets 50% of the neurons to zero during training was also added to prevent overfitting.

To provide a visual representation of this architecture Figure 2 presents a comprehensive diagram showcasing the VGG16 neural network, integrated with the custom fully connected layers described.

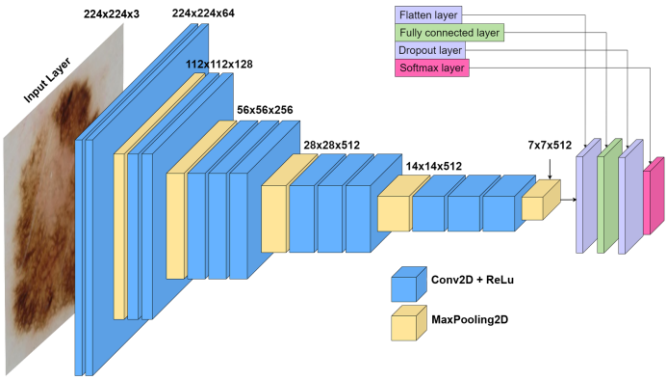


Figure 2. High-level overview of the modified VGG16 network along with the custom fully connected layers added for predictions.

C. Training and Testing

All images were scaled to 224x224x3, and the pixel values were normalized to a range of 0-1 before training the models.

Training was done in two phases:

- Feature Extraction: During this phase, we kept the convolutional base of each model unchanged and focused on training only the additional custom layers we added for classification. We trained these layers for 50 epochs, utilizing data augmentation techniques to generate more training images. The convolution base was frozen, preventing its weights from being updated during training so that the learned features of the model were preserved.

- Fine-tuning: In this phase, we unfroze some of the top layers of the convolutional base of each model and trained them alongside the new layers for 50 epochs. We continued to use data augmentation during this process as well.

D. Data Augmentation

Data augmentation is a technique that applies transformations to the original images to generate new images that can increase the size and diversity of the dataset [18]. In this study, data augmentation was employed due to the limited availability of labeled skin lesion images, which makes it challenging to train accurate and reliable models. To overcome this limitation, the dataset was augmented using various parameters to generate diverse variations of skin lesion images. This augmentation process aimed to enhance the dataset and improve the performance of our models. The parameters used to augment and enhance our dataset are illustrated in Table 1.

Table 1. Parameters used during training phase for all models.

	Parameter	Value
Training	epochs	50 (feature extraction) 50 (fine-tuning)
	batch size	16
Model	optimizer	adam
	loss function	categorical_crossentropy
	metric	validation accuracy
	learning rate	2e-5 (feature extraction) 1e-5 (fine-tuning)
Data augmentation	rotation_range	40
	width_shift_range	0.2
	height_shift_range	0.2
	shear_range	0.2
	zoom_range	0.2
	horizontal_flip	True
	fill_mode	'nearest'

IV. SYSTEM ARCHITECTURE

A Deep Learning virtual machine from Google Cloud equipped with Intel Skylake processor, 16 GB of RAM and a NVIDIA Tesla P100 graphic card with the CUDA 11.0 platform was used to train and build the models.

A. Technologies and libraries employed

TensorFlow: Open-source framework for machine learning and deep learning, developed by Google. Version 2.11 of this framework was used in this study.

Keras: High-level API for building and training deep learning models, integrated with TensorFlow. It offers an easy and intuitive way to define, compile, fit, evaluate, and save models.

Vertex AI: Managed service on Google Cloud that enables users to build, deploy, and manage machine learning models at scale. Google cloud and Vertex AI were employed to store the model artifacts and deploy the model to an endpoint to enable real-time predictions.

React JS: Front-end framework for building user-friendly interfaces using JavaScript. The web-based system was built using this framework.

Gradio: Python library that simplifies the creation of interactive web interfaces for machine learning models with minimal coding effort. The main component for the diagnosis tool was created using Gradio.

Vercel: Provides the necessary infrastructure to build, run, and scale web applications using a variety of languages and frameworks. In this study, Vercel was used to host the web application, making it accessible to users.

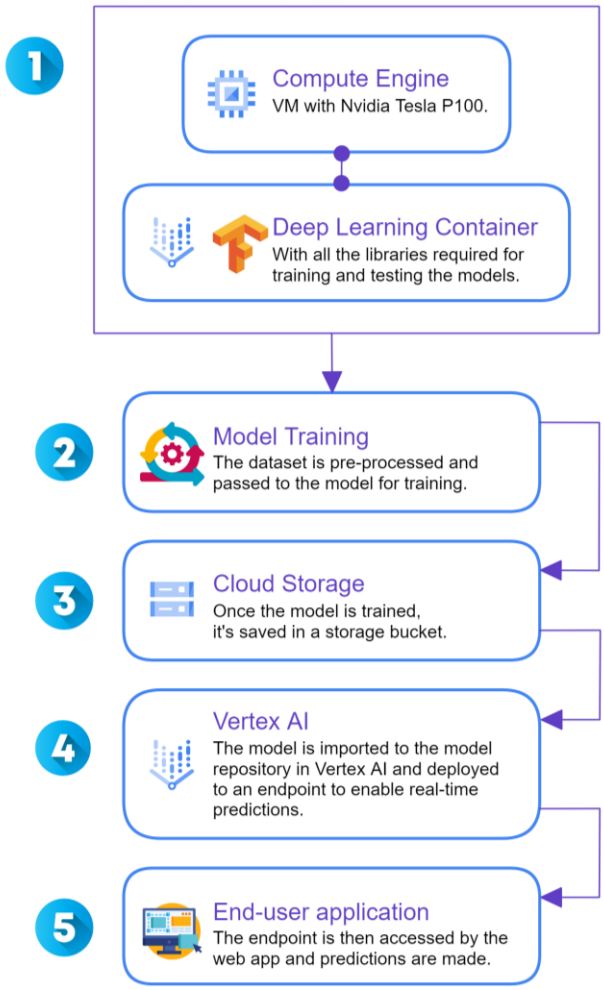


Figure 3. High-level diagram illustrating the functionality of the system.

V. RESULTS

The following section delineates the outcomes derived from the experimental evaluations conducted on the software application that was developed.

Figure 4 displays the results of the training tests, accuracy of the various classification models evaluated.

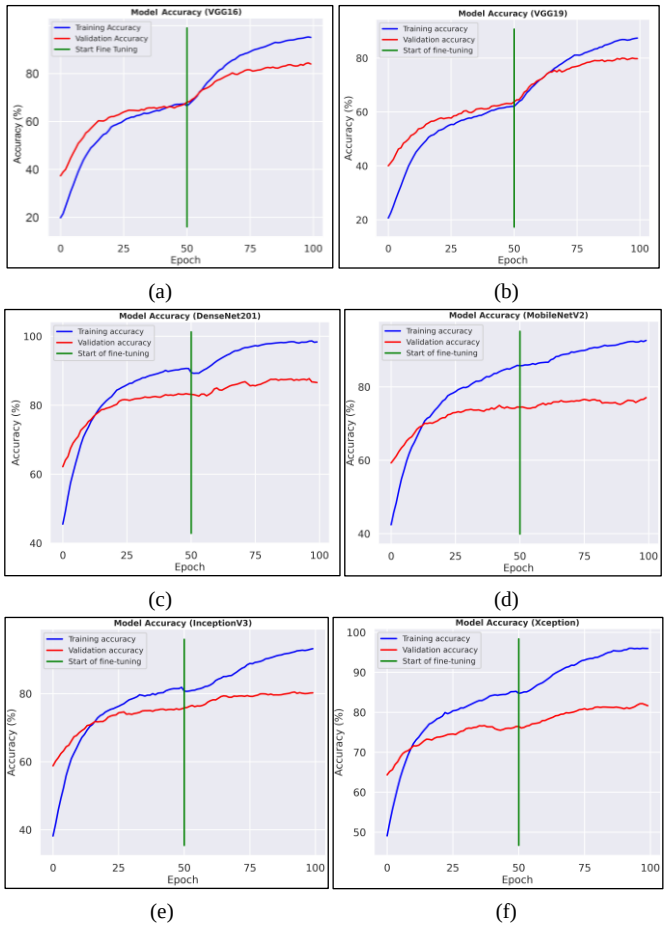


Figure 4: Plots of training and validation accuracy during the phases of feature extraction and fine-tuning: (a) VGG16; (b) VGG19; (c) DenseNet201; (d) MobileNetV2; (e) InceptionV3; (f) Xception.

Table 2. Model performance comparison

Model	Accuracy	Recall	F1 score
VGG16	86.7%	86.7%	86.8%
VGG19	80.0%	80.0%	80.0%
DenseNet201	87.3%	87.3%	87.3%
MobileNetV2	80.6%	80.6%	80.6%
InceptionV3	81.0%	81.0%	80.0%
Xception	81.3%	81.8%	81.4%

Table 3 presents the inference times of the individual models for a single observation, measured using both GPU and CPU on the same machine that was used to train the models. As anticipated, the VGG16 model exhibited the fastest inference time, which, in conjunction with its achieved accuracy of 86.7% as shown in Table 2, was the primary factor for its selection as the model to power our skin lesion diagnosis system. It is worth noting that the DenseNet201 model achieved a slightly higher accuracy of 87.3%, but its inference time was significantly slower in comparison to the other models.

Table 3. Inference time for a single image.

Model	GPU inference time (s)	CPU inference time (s)
VGG16	0.0511	0.1529
VGG19	0.0523	0.1800
DenseNet201	2.5567	0.1450
MobileNetV2	0.0559	0.0551
InceptionV3	0.0788	0.0796
Xception	0.0550	0.1107

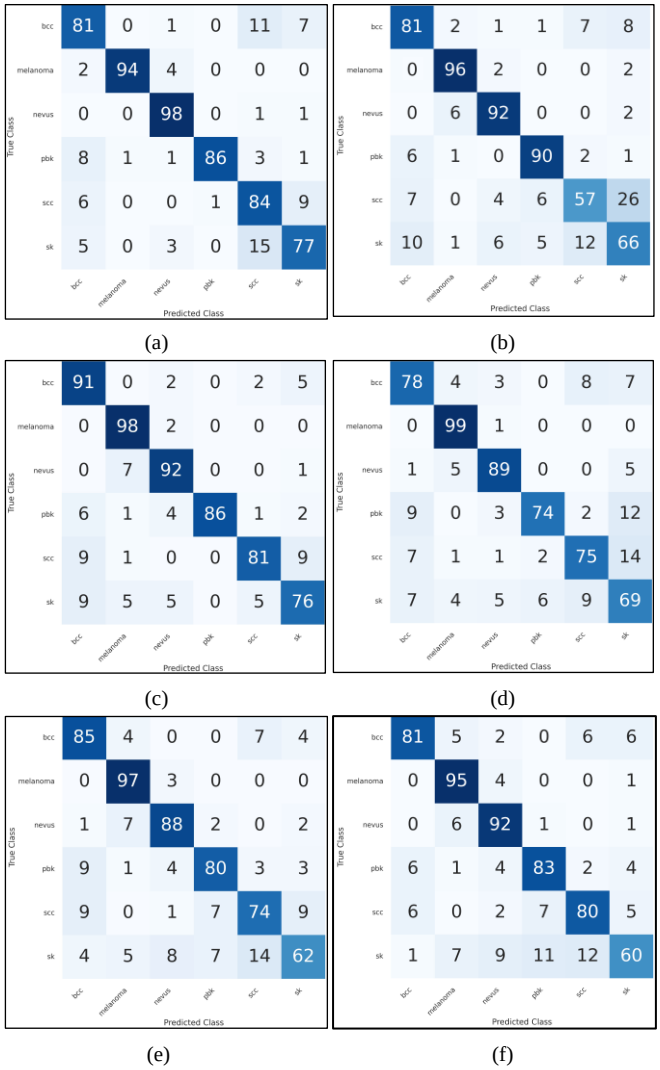


Figure 5. Confusion matrices: (a) VGG16; (b) VGG19; (c) DenseNet201; (d) MobileNetV2; (e) InceptionV3; (f) Xception.

Figure 5 presents the confusion matrix for each of the models studied, with the VGG16 model being the most accurate.

Figure 6 illustrates the fundamental workflow for the prediction process. When a new image is loaded into the system, the user has the option to crop the image and focus on the region of interest (ROI) for the skin lesion. The cropped image is then sent to the VGG16 model endpoint where the image is preprocessed and analyzed. Then the endpoint returns probabilities for each class. These probabilities are subsequently interpreted to determine the corresponding class label, which is then displayed to the user.

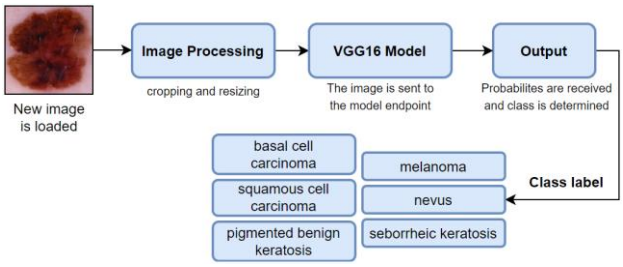


Figure 6. Image processing and prediction workflow of the system powered by VGG16 Model.

Figure 7 displays the interface of the application. The web application is used to upload skin lesion images (left), process them, and make predictions (right).

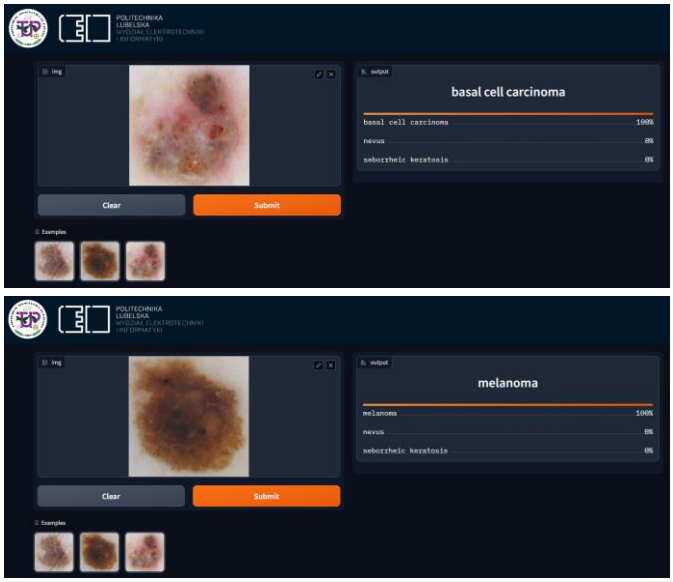


Figure 7. Graphical user interface.

VI. CONCLUSIONS AND FUTURE WORK

The article introduces a web-based system featuring a user-friendly interface designed for the diagnosis of skin lesions utilizing the VGG16 network architecture. This system can identify six distinct types of skin lesions from images uploaded by users, successfully achieving the proposed objective. Six different pre-trained DCNN architectures were compared, and the VGG16 model was identified as the most effective for this task. The models were trained on a dataset of 2400 images, with data augmentation techniques applied to improve performance.

The results of this study demonstrate that it is possible to build an efficient and user-friendly skin lesions diagnosis system, employing machine learning techniques and the latest web technologies. It underscores the advantages of utilizing a web-based platform driven by deep learning models in diagnosing skin lesions, offering potential benefits to both patients and dermatologists.

Future work includes expanding the training set to cover a wider population of dermoscopy images, increasing the number of skin lesions that can be diagnosed by the system, implementing an option for automatic semantic segmentation, possibly powered by [19], and adding some other features to guide users in finding appropriate assistance based on their skin lesion type. These proposed features will undoubtedly enhance the overall performance and ease of use of the system, providing more accurate and reliable results for users.

The web prototype has been developed with the purpose of increasing the likelihood of patients performing proper and fast self-diagnosis. This is crucial at an early stage of the disease to ensure the well-being of the patient.

The final system can be found and tested in the following link: https://dermai-scan.vercel.app/

It is important to highlight that at some point the system could experience interruptions, given that the prediction component has been deactivated to avoid costs associated with the use of cloud services that allow the model to make real-time predictions. To ensure its uninterrupted operation, funding is imperative to keep this project active. Different methods and strategies are currently being evaluated to try to reduce the costs associated with keeping the system active.

REFERENCES

[1] The Skin Cancer Foundation. (2023, March 6). Skin Cancer Facts & Statistics - The Skin Cancer Foundation. https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

[2] Melanoma survival Rates | Melanoma Survival Statistics. (n.d.). American Cancer Society.

https://www.cancer.org/cancer/types/melanoma-skin-cancer/detection-diagnosis-staging/survival-rates-for-melanoma-skin-cancer-by-stage.html

[3] Benedetti, J. (2023, June 21). Diagnostic Tests for Skin Disorders. MSD Manual Professional Edition. https://www.msdmanuals.com/professional/dermatologic-disorders/approach-to-the-dermatologic-patient/diagnostic-tests-for-skin-disorders

[4] Mallick, I., MD. (2022). What is Histopathology? Verywell Health. https://www.verywellhealth.com/histopathology-2252152

[5] Mandal, M. (2023). Introduction to Convolutional Neural Networks (CNN). Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/

[6] Shetty, B., Fernandes, R., Rodrigues, A. P., Chengoden, R., Bhattacharya, S., & Lakshmanna, K. (2022). Skin lesion classification of dermoscopic images using machine learning and CNN. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-22644-9

[7] Ali, M. S., Miah, M. S., Haque, J., Rahman, M. M., & Islam, M. K. (2021). An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. Machine Learning with Applications, 5, 100036. https://doi.org/10.1016/j.mlwa.2021.100036

[8] Kalwa, U., Legner, C., Kong, T., & Pandey, S. (2019). Skin Cancer Diagnostics with an All-Inclusive Smartphone Application. Symmetry, 11(6), 790. https://doi.org/10.3390/sym11060790

[9] Kassem, M. A., Hosny, K. M., & Fouad, M. M. (2020). Skin Lesions Classification Into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning. IEEE Access, 8, 114822–114832. https://doi.org/10.1109/ACCESS.2020.3003890

[10] Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., Halpern, A., Helba, B., Kittler, H., Kose, K., Langer, S., Lioprys, K., Malvehy, J., Musthaq, S., Nanda, J., … Soyer, H. P. (2021). A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Scientific Data, 8(1), 34. https://doi.org/10.1038/s41597-021-00815-z

[11] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[12] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2261–2269. https://doi.org/10.1109/CVPR.2017.243

[13] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[14] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520).

[15] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).

[16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248–255).

[17] Krishnamurthy, B. (2022). An introduction to the RELU activation function. Built In. https://builtin.com/machine-learning/relu-activation-function

[18] Alomar, K., Aysel, H. I., & Cai, X. (2023). Data Augmentation in Classification and Segmentation: A Survey and New Strategies. Journal of Imaging, 9(2), 46. https://doi.org/10.3390/jimaging9020046

[19] A. Kirillov et al., "Segment Anything." arXiv, 2023. doi: 10.48550/ARXIV.2304.02643.