Multi-classification of skin lesion using deep learning techniques

Amel Ykhlef[1] and Menaouer Brahami[2][0000−0003−0045−9797]

[1] Computer Systems Engineering, LABAB Laboratory,National Polytechnic School of Oran,Algeria
amel.ykhlef@enp-oran.dz
[2] Computer Systems Engineering, LABAB Laboratory,National Polytechnic School of Oran,Algeria
menaouer.brahami@enp-oran.dz

Abstract. Skin cancer is one of the most dangerous diseases in the world and accounts for the highest number of deaths each year. According to the Skin Cancer Foundation, the number of newly diagnosed melanoma cases in the United States is expected to increase by 7.3% in 2024, and the number of melanoma deaths is expected to increase by 3.8%, This is why artificial intelligence techniques have been used to solve skin cancer problems, using deep learning techniques for classification, detection and prediction. In this work, we present a deep learning technique for object detection, known as YOLO (You Only Look Once) as a first step rather than using convolutional neural network (CNN) for multi-classification. The results showed that the model achieved accuracy of 94% for testing data and 98 % for training data. This methodology shows that the use of object detection techniques can improve classification performance.

Keywords: Skin Lesion · Deep Learning · YOLO v8 · CNN · Multi-classification.

1 Introduction

Skin diseases, particularly skin cancer, present significant challenges in early diagnosis due to the visual similarity between the different types of lesions and the challenge of their classification. Modern studies have shown that Artificial Intelligence (AI) has been used in several domains, one of these domains is medical field. According to [8], AI has drawn tremendous attention for its potential utility in every facet of human activity, including healthcare. Skin cancer is one of the world's most dangerous diseases due to the high mortality percent, many researchers followed this problem by using various AI techniques to improve the efficacy of AI models in medical data. The use of several deep learning techniques helps in the early detection of disease. As a first step, skin lesion localization is done by several segmentation methods, among which we used Yolo since, according to previous research [1, 2] it has achieved good results in object detection. Moreover, the YOLO v8 algorithm of computer vision is utilized in various Artificial Intelligence applications to identify and localize items inside an image and

allow a machine to learn on its own. After object detection, we classify lesions into seven classes: Akiec (Actinic Keratosis), BCC (Basal Cell Carcinoma), BKL (Benign Keratosis), DF (Dermatofibroma), MEL (Melanoma), NV (Melanocytic Nevi), VASC (Vascular) using a convolutional neural network.
This paper introduces an approach for object detection and multi-classification of skin lesion images, proposing the following objectives:

- A YOLO v8 model for identifying the region of interest (ROI) in skin lesion images.
- A multi-classification model utilizing a convolutional neural network (CNN).

The paper is organized as follows: section 2 describes the previous research related to skin lesion classification and segmentation, section 3 presents the methodology used in this work with detailed explanation to classify the skin lesion images, section 4 discussed the experimental setup and results and section 5 for conclusion.

2 Related Work

Many studies have mentioned the impact of AI techniques and technologies in skin lesion detection and classification. For example, Wang et al. [1] proposed a Residual Attention Yolo v4 (RA-YOLOv4) (residual backbone, a lightweight Pyramid Attention with Simple Network Backbones (PANet-s), and a yolo head) model, which was evaluated on the ISIC-skin dataset 2017 and 2018 to detect melanoma, benign nevus, and benign keratosis. The model can locate and classify the three types of lesions. Elshahawy et al. [2] proposed a hybrid model based on Yolo v5 and ResNet network to detect melanoma on a publicly available dataset (HAM10000) with an accuracy of 99.5%. Besides that, Oukil et al. [3] proposed an algorithm based on feature extraction using three classifiers (k-nearest neighbours, support vector machine, and artificial network) for the color and texture of skin lesions using the PH2 dataset and a segmentation technique based on k-means used to localize the lesion. This method found good results by reaching the accuracy of 99.51% with an SVM classifier and RGB color. In another study that was done by Khan et al [4], which proposed the use of a pretrained deep neural network (DarkNet 19) and fine- tuned the parameters of the third convolutional layer to classify lesions taken from HAM 10000, ISBI2018, and ISBI 2019 datasets. In a study conducted by Anupama et al. [5] present a model (deep learning with image segmentation based on an Evolutionary algorithm. Orhan and Yavsan [6] proposed an artificial intelligence model for skin cancer detection using deep learning algorithms such as MobileNet, AlexNet, ResNet, VGG16, and VGG19. the MobileNet model showed the highest performance. The model has achieved an accuracy of 84.94%. Another research by Ykhlef et al. [9] mentioned in their works the difference between YOLO v8 and YOLO v9 on the detection of objects in skin images and both achieved good results with mAP equal to 98,7 %. Based on previous studies and the significant progress that has been made in this field, we proposed an approach that allows us to detect and classify skin lesion images of HAM 10000 dataset.

3 Approach Proposed

As per our objective and motivations, this study is associated with some background ideas and research efforts as shown in Figure 1. We employed the YOLO v8 object detection model to isolate the region of interest (ROI), followed by resizing the images for input into a classification model. A convolutional neural network (CNN) was utilized to classify the ROI into seven skin lesion categories: melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AKIEC), benign keratosis (BKL), dermatofibroma (DF), and vascular lesion (VASC). The proposed methodology is depicted in Figure 1:

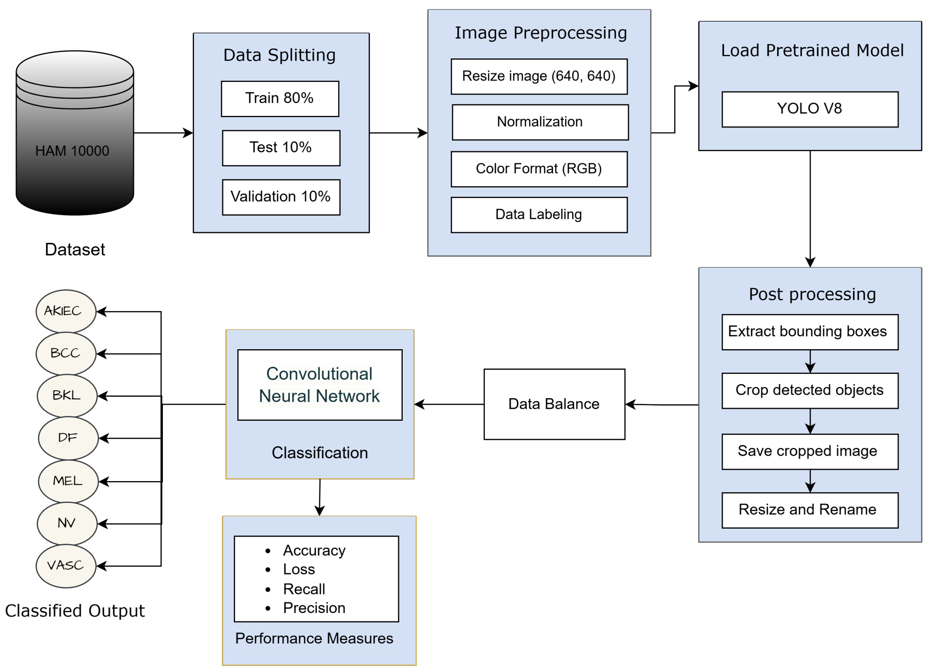

Fig. 1. Proposed Approach.

3.1 Dataset Description

The data used for this proposed approach are a publicly published dataset that contains 10015 images, classified into 7 classes: melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AKIEC), benign keratosis (BKL), dermatofibroma (DF), and vascular lesion (VASC). as shown in Figure 2:

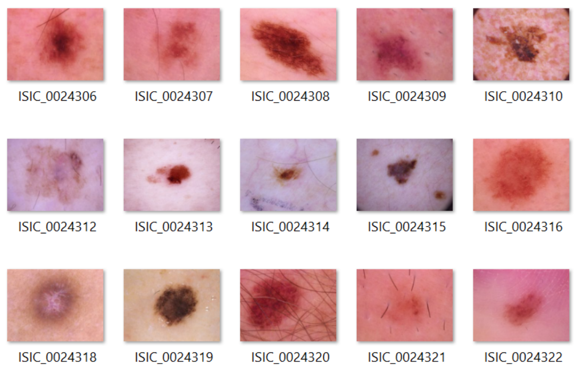

Fig. 2. Dermoscopic skin lesion image (HAM 10000)

The following Table1 shows the different classes in the HAM 10000 dataset. The first column shows the class name, the second column shows the full expression of the class, and the third column shows the number of images belonging to each of the 7 classes.

Table 1. Classes of HAM 10000 dataset.

Class Labels	Full expression	Number of images
Akiec	Actinic Keratosis	327
BCC	Basal Cell Carcinoma	514
BKL	Benign Keratosis	1099
DF	Dermatofibroma	115
MEL	Melanoma	1113
NV	Melanocytic Nevi	6705
VASC	Vascular Lesion	142

3.2 Data Splitting

Data splitting before using YOLO v8 refers to dividing data into training, validation, and testing. This division ensures that the model learns from the training set, while its performance is monitored and validated on unseen data (validation set) and finally assessed on completely independent data (test set). Proper data splitting helps prevent overfitting, enables reliable tuning of model parameters, and provides an unbiased evaluation of how well the model generalizes to new images. In object detection tasks with YOLO v8, it is important that these splits maintain a representative distribution of object classes and annotations to ensure balanced learning and fair assessment.

3.3 Data Preprocessing

The most important step in preparing the image for training deep learning models is to prepare four main tasks:

- Resize image: images are resized to uniform dimensions (640, 640 PIXELS).
- Normalization: The raw pixel intensity values of an image are usually between 0 and 255. Normalizing reduces them to a lower range, usually 0 to 1, by dividing all pixels by 255. Normalization stabilizes and speeds up model training by making input data more standardized.
- Color format RGB: The RGB color format represents images using three color channels-red, green, and blue, where each pixel color is created by combining different intensities of these three colors. It is a common way to display and process color images in digital systems.
- Data Labeling: image labeling provides bounding box coordinates and class labels for all objects so the model can be trained to identify and find objects appropriately in object detection tasks.

3.4 Object detection

Computer vision technology called object detection, is used to find and identify elements in an image or video. For effective classification, we used the pre-trained Yolo v8 models to detect the region of interest in the image of the skin lesion. After training the model, YOLO v8 achieved high accuracy. The Figure 3 illustrates the architecture of the YOLO v8 model used in our approach.

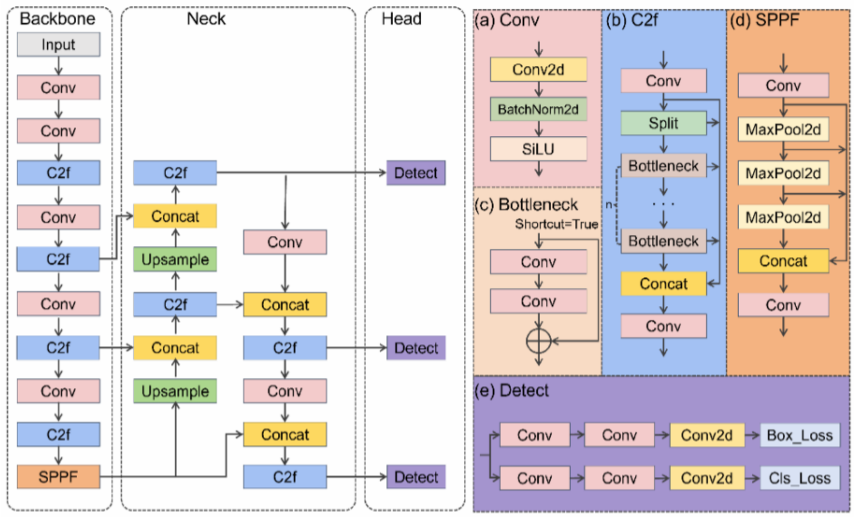

Fig. 3. The architecture of Yolo v8.

The following Figure 4 illustrates how the YOLO v8 model detects regions of interest within an input image by drawing bounding boxes around identified objects. Each bounding box is accompanied by an accuracy value (confidence score) that indicates the model's certainty in the detection. The YOLO v8 model can predict the region of interest (ROI) with high performance that varies between 90% and 100%.

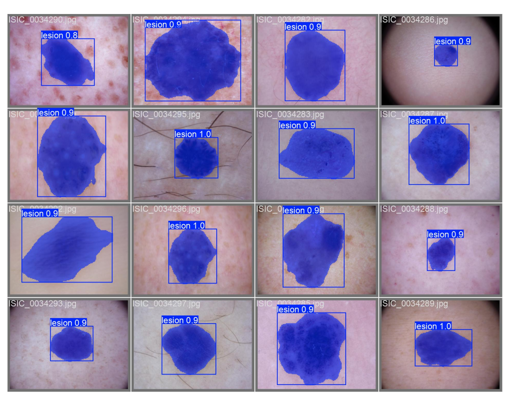

Fig. 4. Lesion detection using YOLO8.

3.5 Post processing

After using YOLOv8, the detected bounding boxes are used to crop the corresponding regions from the original images. These cropped images are saved individually for further analysis and classification. This post-processing step enables focused examination of each detected object and facilitates classification, renaming the cropped image to help identify each sample, and resizing the images to a uniform dimension (224,224) pixels, ensuring that all inputs to the CNN models have the same size.

3.6 Data Balance

To address class imbalance on our dataset, we applied the resampling technique and this helped in increasing the minority class samples, this method helped in creating a more balanced dataset. For definition, Resampling is a statistical method that involves repeatedly drawing samples from a data set to assess the variability of a statistic without relying on strict distributional assumptions[10].

3.7 Convolutional Neural Network

Convolutional neural network (CNN) was used to classify the cropped region of interest (ROI), the output of the proposed model is 7 classes: melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AKIEC), benign keratosis (BKL), dermatofibroma (DF), and vascular lesion (VASC). The layers of the Convolutional Neural Network (CNN) used:

- Conv 2D: the fundamental building blocks of convolutional neural network designed to process images.
- MaxPooling 2D layer: takes the max values of feature map, reduces height and width on feature map while retaining the important information.
- Dropout: is a regularization technique used to prevent overfitting.
- Flatten: converts the output of conv2D or maxpooling from multi-dimentional to 1D vector.
- Dense: known as fully connected layer, and it is the last step on CNN used for classification. Hierarchical feature learning, where the early layers detect simple features, and the deeper layers combine them into complex representations.

The detailed architecture of the Convolutional Neural Network (CNN) utilized in this study is presented in Figure 5. This figure demonstrates the sequential organization of layers.

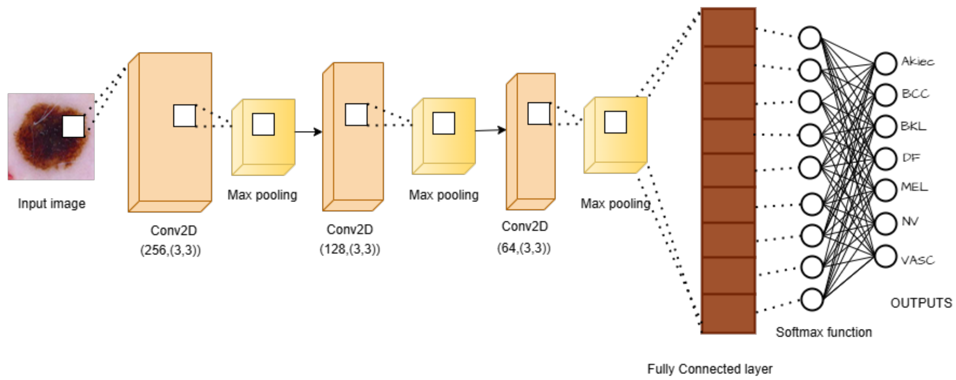

Fig. 5. CNN architecture.

The Table 2 presents the architectural components of CNN model used in this study, including convolutional layers to process images, pooling layers to reduce dimensionality, fully connected layers for classification, dropout layers for regularization and softmax function was used for multi-classification.

Table 2. Representing the layers of CNN model.

Layer type	Parameters	Output shape
Conv2D	256 filters, 3x3 kernel	(30, 30, 256)
MaxPooling2D	pool size=(2, 2)	(15, 15, 256)
Dropout	rate=0.5	(15, 15, 256)
Conv2D	128 filters, 3x3 kernel	(15, 15, 128)
MaxPooling2D	pool size=(2, 2)	(6, 6, 128)
Dropout	rate=0.5	(6, 6, 128)
Conv2D	64 filters, 3x3 kernel	(6, 6, 64)
MaxPooling2D	pool size=(2, 2)	(4, 4, 64)
Dropout	rate=0.5	(4, 4, 64)
Flatten		(4 * 4 * 64,)
Dense	32 units	(32,)
Dense	7 units, softmax activation	(7,)

4 Results and discussion

We evaluate the performance of the proposed CNN model based YOLO v8 for multi-classification using several metrics. Confusion Matrix shown in Figure 6 represents the predicted labels against the actual true labels the matrix prediction is divided into four: True Positive(TP), True Negative(TN), False Positive(FP) and False Negative(FN). The confusion matrix in Figure 6 shows that class zero (Akiec) and class three (DF) and class six (VASC) are well classified so the model predicts all the true labels of these three classes, focusing on the classification of NV class 412 images were well predicted represents True Positive (TP) while False Negative (FN) 48 were predicted as melanoma, 23 images as BKL, 6 as BCC, 4 as VASC and 2 as Akiec.

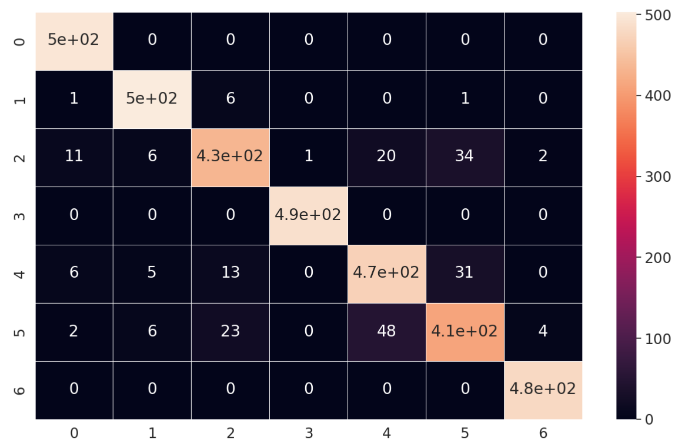

Fig. 6. Confusion Matrix.

The provided Table 3 summarizes the Precision, Recall, F1-score and support of each class in the HAM 10000, the table shows that the minimum value of precision is 86% of the NV class and the highest precision was done by DF class which means that the predicted samples on this class were correct 100% , the recall for classes Akiec, DF, BCC and VASC showed that the model were correctly identified 100% and 98% of all true class,the highest value of F1-score was done by class DF which reflect a good balance between precision and recall,

Table 3. Performance comparison of different classes of the dataset

	Precision	Recall	F1-score	Support
Akiec	0.96	1.00	0.98	496
BCC	0.97	0.98	0.98	511
BKL	0.91	0.85	0.88	508
DF	1.00	1.00	1.00	489
MEL	0.87	0.90	0.88	524
NV	0.86	0.83	0.85	494
VASC	0.99	1.00	0.99	478

To assess the effectiveness of the CNN based on YOLO v8 for multiclassification accuracy graph, it is necessary to show how well the model correctly classifies HAM 10000 into seven classes, the model achieved a high accuracy with 98% for training and 94% for testing against the number of training epoch 100, this value represents how the predicted value increased while using the model proposed, the graph below illustrates the model's accuracy over 100 epoch see Figure 7

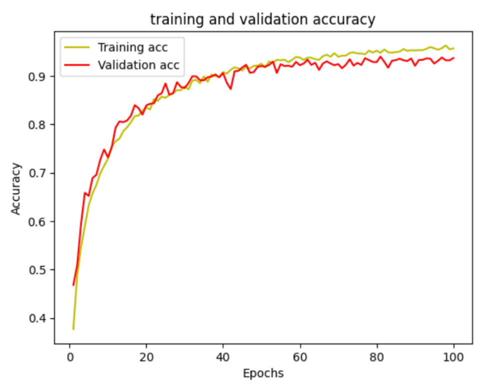

Fig. 7. Accuracy graph (a) training graph (b) testing graph.

The graph below shows the training and testing loss over epochs, indicating how the model's prediction errors decrease during learning see Figure 8.

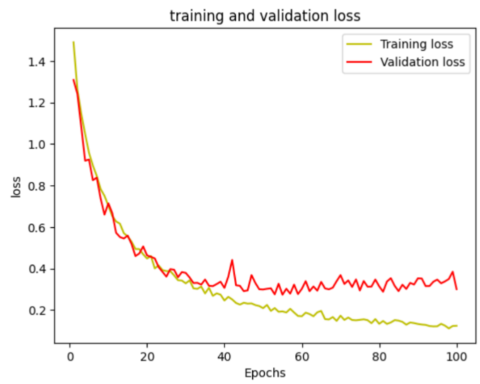

Fig. 8. Loss graph (a) training graph (b) testing graph .

The recall graph shows how well the model recognizes relevant examples from all the real positive sources. A higher recall value in the graph signifies better performance in detecting related cases, which is especially important in applications where missing positive instances carries a high cost see Figure 9.

Fig. 9. Recall graph (a) training graph (b) testing graph .

In summary, these metrics offer valuable insights into the model's reliability, precision, and areas for improvement.

5 Practical and Managerial Implications of the findings

This study accelerates the diagnostic process and simplifies the analysis of medical data by enhancing the ability to distinguish between skin cancer images and other types of cancer. As a result, it contributes to the reduction of mortality rates. Healthcare professionals can leverage our method to focus their expertise more effectively and develop improved screening protocols, thus improving the overall quality and efficiency of patient care.

6 Conclusion

Many disorders connected to the skin can only be diagnosed with the use of image analysis. Early detection of a skin lesion and the results of the initial smears can result in an instant diagnosis and the start of therapy. Deep learning has been used to solve problems of image processing in recent years. This study focused on the multi-classification of the HAM 10000 dataset into seven categories: actinic keratosis (Akiec), basal cell carcinoma (BCC), benign keratosis (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic Nevi (NV), and vascular lesion (VASC). YOLO v8 was utilized for object detection to identify and crop the region of interest (ROI) from the HAM 10000 images. The cropped ROI was then used to train a convolutional neural network (CNN) model, which demonstrated strong performance achieving 94% accuracy on testing data and 98% accuracy on training data. The proposed approach exhibited high effectiveness in classifying skin lesion images with excellent precision.

References

1. Wang, S., Luo, J., Zhou, Q., Ren, X., Zhang, Y.: A Differential Diagnose Method for Dermoscopy Images. In: Proceedings of the 2023 International Conference on Advanced Computational Intelligence (ICACI), pp. 1–8 (2023). doi:10.1109/ICACI58115.2023.10146178.
2. Elshahawy, M., Elnemr, A., Oproescu, M., Schiopu, A.G., Elgarayhi, A., Elmogy, M.M., Sallah, M.: Early Melanoma Detection Based on a Hybrid YOLOv5 and ResNet Technique. Diagnostics 13(17), 2804 (2023). doi:10.3390/diagnostics13172804.
3. Oukil, S., Kasmi, R., Mokrani, K., García-Zapirain, B., (2021), Automatic segmentation and melanoma detection based on color and texture features in dermoscopic images, doi:10.1111/srt.13111.
4. Khan, M.A., Akram, T., Sharif, M., Kadry, S., Nam, Y.: Computer Decision Support System for Skin Cancer Localization and Classification. Computers, Materials & Continua 68(1), 1041–1064 (2021). doi:10.32604/cmc.2021.01630.
5. Anupama, C. S. S., Natrayan, L., Laxmi Lydia, E., Sait, A. R. W., Escorcia-Gutierrez, J., Gamarra, M., & Mansour, R. F., (2022), Deep Learning with Backtracking Search Optimization Based Skin Lesion Diagnosis Model, DOI: DOI:10.32604/cmc.2022.018396.
6. Orhan, H., & Yavşan, E. (2023). Artificial intelligence-assisted detection model for melanoma diagnosis using deep learning techniques. Mathematical Modelling and Numerical Simulation with Applications, 3(2), 159–169.
7. Ykhlef A, Labri NS, Brahami M. Blood Product prediction using Supervised Machine Learning. In: The first international Conference on Advances in Electronics, Control and Computer Technologies(ICAECCT'23), October 25-26, Mascara, Algeria. 2023.

8. Ykhlef, A., Labri, N. S., & Brahami, M. (2024). Supervised learning techniques for blood product prediction in patients with hematologic diseases: A multi-center study in Western Algeria. International Journal of Information Technology, 16(5), 1–26. https://doi.org/10.1007/s41870-024-01928-5.
9. Ykhlef, A., Brahami, M. (2024). Detecting lesion on dermoscopic images using YOLO algorithms. In A. Khelassi (Ed.), Abstract book International Congress on Health Science and Medical Technologies 2024. Knowledge Kingdom Publishing. https://doi.org/10.26415/978-9931-9446-8-3
10. Efron, B., & Tibshirani, R. J. (1993). An Introduction to the Bootstrap. Chapman & Hall/CRC.