

Enhanced Model for Classifying Skin Diseases Using YOLO Technique

Mohamed Saied El-Sayed Amer

Canadian International College, New Cairo, Egypt

DOI: <https://doi.org/10.51244/IJRSI.2025.121500058P>

Received: 02 April 2025; Accepted: 04 April 2025; Published: 03 May 2025

ABSTRACT

One of the most prevalent disorders is skin disease. Skin disorders are difficult to classify because of their complex classifications, early-stage symptoms that are quite similar, and highly imbalanced lesion samples. Simultaneously, given a small amount of data, a single trustworthy convolutional neural network model has poor generalization capacity, insufficient feature extraction capability, and low classification accuracy. Thus, based on model fusion, we suggested a classification model based on YOLO for the categorization of skin diseases in this research. A computer vision model in the You Only Look Once (YOLO) family is called YOLO. YOLO is frequently utilized for object detection. YOLO is available in four primary variants, with increasing accuracy rates: small (s), medium (m), large (l), and extra large (x). The training time for each version varies as well.

Keywords: skin disease classification, object detection, DarkNet, YOLO, skin disorder

INTRODUCTION

A serious global public health issue that impacts a lot of people is skin disease (Karimkhani et al., 2017). Many skin illnesses have different symptoms, and it takes time for those symptoms to change. The majority of individuals frequently ignore changes in their skin symptoms, which can have serious repercussions like permanent skin damage and an increased risk of skin cancer (Leiter, U. et al. 2014). It is also difficult for the average person to identify the type of skin condition with the naked eye. Furthermore, morbidity and death from skin cancer can be reduced with early therapy (Baumann, B.C. et al., 2020).

Furthermore, deep learning has quickly become the recommended technique for medical image analysis as a result of its rapid progress (Litjens, G. et al., 2017; Abdulrahman, A.A. et al., 2020). Although, deep learning exhibits superior generalization ability and increased robustness when compared to other classification techniques (Rashid, T. et al., 2022). Convolutional neural networks are currently among the most popular and representative models for deep learning (Liu, W. et al. 2017; Pouyanfar, S. et al., 2018). Medical image classification has advanced significantly, and it is now widely employed in various facets of medical image analysis (Ker, J. et al., 2017; Anwar, S.M. et al., 2018).

With this combination, the sensitivity score on the ISIC2017 (Codella, N.C. et al., 2017) dataset was enhanced to 91.6% when compared to the baseline model. Other than that, it outperformed the baseline model by 4.7% with an accuracy of 93.7% on the HAM10000 (Tschandl, P. et al., 2018) dataset. FixCaps is a capsule network approach that was proposed by Lan. et al. (Lan, Z. et al., 2022). With a wider receptive domain, it is an enhanced convolutional neural network model built on CapsNets (Sabour, S. et al., 2017). It functions by applying a large, high-performance kernel at the bottom convolutional layer, with a kernel size of up to 31×31 . Simultaneously, an attention mechanism was implemented to mitigate the loss of spatial information resulting from convolution and pooling. The HAM10000 dataset demonstrated an accuracy of 96.49% and a f1-score of 86.36%.

Both the FixCaps and IRV2-SA models perform brilliantly in terms of classification accuracy. They perform unsatisfactorily in classifications with little individual sample data, and they do not meet all other standards for

classification performance evaluation. Improving their classification accuracy is hampered by the large imbalance of lesion samples and the paucity of imaging data available for skin illnesses.

By simultaneously identifying and classifying skin lesions in a single, quick pass, YOLO tackles the difficulties in skin disease classification in this study, allowing for real-time analysis that is essential for clinical and mobile applications. Low inter-class variance is lessened by its efficient use of bounding boxes to isolate the region of concern, which lessens the effect of background noise and increases emphasis on the lesion itself. YOLO's end-to-end architecture streamlines the pipeline by doing away with the requirement for independent segmentation or cropping phases, and its compatibility with data augmentation increases its resilience against high intra-class variance and class imbalance.

However, by integrating detection and classification into a single, effective framework that can precisely pinpoint lesions and function in real time, YOLO addresses the difficulties associated with classifying skin diseases. It helps isolate the region of concern by drawing bounding boxes around impacted areas, which lessens confusion brought on by background noise or similar-looking situations. Data augmentation strategies also help YOLO manage patient appearance variances and enhance generalization. YOLO is ideally suited for applications like teledermatology or mobile diagnostics where accuracy and speed are crucial because it concentrates on lesion-specific regions and provides quick inference.

Furthermore, the complex classifications of skin illnesses and the early parallels in their symptoms complicate the model classification. At the same time, a single reliable network model tested with limited data has a weak generalization capacity and insufficient feature extraction capability.

Accurate classification with great precision is still a challenge. Data augmentation, or increasing the model's ability to extract features, is a popular study strategy for resolving the problems of small sample sizes and class imbalance.

Related work:

Many CNN models have been investigated for the classification of skin diseases, and some of these models have demonstrated excellent classification performance. The pertinent published work of a few researchers in the field of image classification for skin diseases is summarized below.

Several reputable multi-class CNN models have been proposed by researchers. Mobiny et al. (Mobiny, A. et al., 2019) described The Bayesian DenseNet-169, an approximate risk-aware deep Bayesian model, generates an assessment of model uncertainty without requiring the addition of new parameters or a major modification to the network topology. It improved the underlying DenseNet169 (Huang, G. et al., 2017) model's classification accuracy on the HAM10000 dataset from 81.35% to 83.59%. An interpretability-based CNN model is put forth by Wang et al. (Wang, S.; et al., 2021). In order to diagnose skin lesions, this multi-class classification model uses patient metadata and pictures of the lesions as input. Its accuracy and sensitivity on the HAM10000 dataset were 95.1% and 83.5%, respectively.

Allugunti et al. (Allugunti, V.R. et al., 2022) developed a multi-class CNN model to diagnose skin cancer. A differentiation between lesion malignant, superficial spreading, and nodular melanoma is made in the suggested model. This makes it possible to identify the virus early and to begin the treatment and isolation process as soon as possible to prevent the infection from spreading. The Xception (Chollet, F. 2017) model was altered by Anand et al. (Anand, V. et al., 2022) by include layers like a dropout layer, two thick layers, and a pooling layer. Seven kinds of skin diseases were added to the original fully connected (FC) layer in a new FC layer. On the HAM10000 dataset, its classification accuracy was 96.40%.

Using ensemble learning to increase the model's classification accuracy is another useful strategy. For the purpose of classifying skin lesions, Thurnhofer-Hemsi et al. (Thurnhofer-Hemsi, K. et al., 2021) suggested an ensemble made up of enhanced CNNs in conjunction with a regularly spaced test-time-shifting method. It uses a shift technique to build up several test input images, which are then delivered to each classifier in the ensemble. Finally, it combines all of the outputs for classification. On the HAM10000 dataset, its classification accuracy was 83.6%.

An attention module can be added to a model to improve its capacity for extracting features, which will improve the model's performance in classification. When Karthik et al. (Karthik, R. et al., 2022) substituted an Efficient Channel Attention (Wang, Q. et al., 2020) block for the conventional Squeeze-and-Excite (Hu, J.; Shen, L. et al., 2018) block in the EfficientNetV2 (Tan, M.; Le, Q. 2021) model, the overall number of training parameters decreased dramatically. In four different categories of skin disease datasets—acne, actinic keratosis, melanoma, and psoriasis—the test accuracy of the model was 84.70%.

The accuracy of image categorization can be improved by using image processing techniques such as segmentation, equalization, enhancement, and conversion. An enhanced data augmentation model is suggested by Abayomi-Alli et al. (Abayomi-Alli, O.O. et al., 2021) for the successful identification of melanoma skin cancer. To develop synthetic melanoma pictures, the method relied on oversampling data embedded in a nonlinear low-dimensional manifold. On the PH2 (Mendonça, T. et al., 2013) dataset, it obtained accuracy, sensitivity, specificity, and f1-score of 92.18%, 80.77%, 95.1%, and 80.84%, respectively.

Hoang et al.'s (Hoang, L.; Lee, S.-H. et al., 2022) novel method of categorizing skin lesions makes use of a new segmentation strategy and wide-ShuffleNet. It initially separates the lesion from the surrounding area by computing an entropy-based weighted sum first-order cumulative moment (EW-FCM) of the skin image. Following segmentation, a unique deep learning structure known as wide-ShuffleNet is used to classify the data. It achieved specificity, sensitivity, precision, f1-score, accuracy, and 76.15%, 72.61%, and 84.80%, respectively, on the HAM10000 dataset.

Malibari et al. (Malibari, A.A. et al., 2022) proposed an ideal Deep-Neural-Network-Driven Computer-Aided Diagnosis Model for their skin cancer detection and classification model. The model mostly uses a U-Net segmentation technique after a Wiener-filtering-based pre-processing step. The model's highest accuracy was 99.90%.

Nawaz et al. (Nawaz, M. et al., 2022) presented an enhanced Deep Learning-based technique, the DenseNet77-based UNET model. Their tests proved the model's dependability and its capacity to recognize skin lesions of various hues and sizes. On the ISIC2017 [11] and ISIC2018 (Codella, N. et al., 2019) datasets, it achieved an accuracy of 99.21% and 99.51%, respectively.

Therefore, we suggested a YOLO model for skin disease classification by synthesizing the relevant work done by these academics in the field of skin disease picture classification.

METHODOLOGY

First, using our dataset, which is dominated by kaggle, we trained and evaluated the classification performance of a basic YOLO model. This dataset was typical in that it had very imbalanced categories and a tiny sample size. Following the investigation, it was discovered that the YOLO model, which was the best-performing model, had an excellent classification performance with accuracy rates of 95.3%.

The dataset description:

The dataset contains 878 images for skin diseases which are Actinic keratosis, Atopic Dermatitis, Benign keratosis, Dermatofibroma, Melanocytic nevus, Melanoma, Squamous cell carcinoma, Tinea Ringworm Candidiasis and Vascular lesion. This dataset is splitted into two categories which are “train” that contains 697 images and “test” that contains 181 images. Each image is labeled with the corresponding disease and prepared for the model training.

The model:

The YOLO model generates features for object detection based on input images. These properties are then used by a prediction system to draw boxes around objects and identify the classes to which they belong. In this stage, the photos are categorized using the YOLO based on the traits that were identified.

The first object detector to link the process of class label prediction with bounding box prediction in an end-to-end differentiable network was the YOLO model. Figure 1 shows three primary components for the YOLO network which are:

- *Backbone*: An image feature-aggregating and feature-forming convolutional neural network at various granularities.
- *Neck*: An arrangement of layers that blends and combines picture characteristics before forwarding them to prediction.
- *Head*: Uses characteristics from the neck to determine classes and boxes.

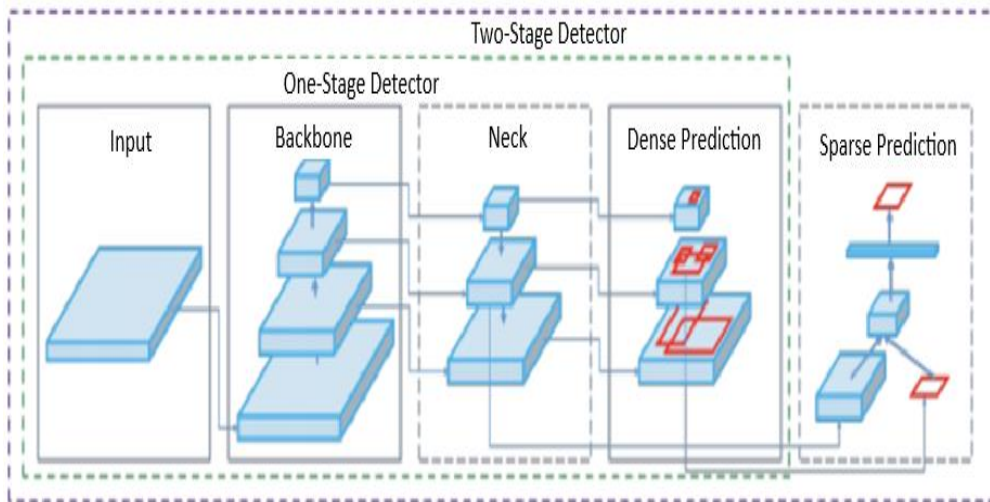


Figure 1: YOLO components (Bochkovskiy et al., 2020)

Having stated that, there are numerous ways to combine various architectures at each significant component. YOLO's primary contribution is to incorporate advances from other fields of computer vision and demonstrate how, taken together, they enhance YOLO object detection.

Though they are frequently less mentioned, the steps used to train a model are just as crucial to the overall effectiveness of an object recognition system as any other component. Let's discuss the two primary YOLO training processes:

- **Data augmentation**: involves modifying the initial training data to expose the model to a larger variety of semantic variation than what would be found in the training set alone.
- **Loss Calculations**: YOLO uses the GIoU, obj, and class losses functions to compute a total loss function. The goal of mean average precision can be maximized by carefully crafting these functions.

YOLO's translation of the Darknet research framework to the PyTorch framework is its most significant contribution. With its fine-grained control over the processes encoded into the network, the Darknet framework is mostly developed in C. The control over the lower level language is helpful for research in many ways, but it can also slow down the translation of new research findings because each new addition requires the writing of a specific gradient calculation. It is no small task to translate the training protocols from DarkNet to PyTorch in YOLO and beyond.

Experiments and results:

The research has been implemented using Python. Around 878 images are taken from a dataset for dermatologist for each category of disease. These images were then preprocessed through image labeling using image labeling tool. Figure 2 shows a sample of the dataset for 7 different skin diseases:



Figure 2: Dataset sample (Riya Eliza Shaju, 2023)

Every image in the sample has a label corresponding to the name of the disease, and the disease name corresponds to the class name that is used to represent the disease during training and detection.

YOLO runs training data through a data loader, which enhances data online, with each training batch. Three types of augmentations are made by the data loader:

- **Scaling:** refers to how the YOLO (You Only Look Once) architecture is adapted and optimized for different levels of performance, typically balancing accuracy, speed, and model size.
- **Color space adjustments:** refer to preprocessing techniques used to enhance the model's ability to detect and classify objects by modifying how color information is represented in the input images.
- **Mosaic augmentation:** is a powerful data augmentation technique used in YOLO to improve model robustness and detection performance

In the context of the popular COCO object recognition benchmark, mosaic augmentation is especially useful in solving the well-known "small object problem". This challenge has to do with how smaller items are perceived more inaccurately than larger ones.

The experiment conducted on the dataset using operating system Windows 11, which was configured with Core-i7 CPU and 4GB NVIDIA 16GB GPUs. The experiment run on 100 epochs on the labeled dataset. Figure 3 illustrate the training steps for the experiment which contains the losses and instances of the training process:

Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:55:00:00, 8.27s/it]
36/99	0G	0.0414	0.01884	0.01959	11	640	100%	14/14	[01:55:00:00, 8.27s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.77s/it]
	all	213	280	0.615	0.702	0.665	0.397		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:54:00:00, 8.21s/it]
37/99	0G	0.0382	0.01938	0.01803	16	640	100%	14/14	[01:54:00:00, 8.21s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.81s/it]
	all	213	280	0.597	0.769	0.74	0.445		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:55:00:00, 8.27s/it]
38/99	0G	0.0383	0.01951	0.01923	22	640	100%	14/14	[01:55:00:00, 8.27s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.78s/it]
	all	213	280	0.564	0.759	0.711	0.41		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:54:00:00, 8.21s/it]
39/99	0G	0.04113	0.02122	0.0237	22	640	100%	14/14	[01:54:00:00, 8.21s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.80s/it]
	all	213	280	0.459	0.562	0.52	0.31		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:55:00:00, 8.21s/it]
40/99	0G	0.03919	0.01832	0.02132	15	640	100%	14/14	[01:55:00:00, 8.21s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.75s/it]
	all	213	280	0.539	0.668	0.621	0.382		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:55:00:00, 8.26s/it]
41/99	0G	0.03965	0.02039	0.01671	26	640	100%	14/14	[01:55:00:00, 8.26s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:34:00:00, 4.89s/it]
	all	213	280	0.555	0.742	0.721	0.442		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:53:00:00, 8.09s/it]
42/99	0G	0.0391	0.02124	0.0173	20	640	100%	14/14	[01:53:00:00, 8.09s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.81s/it]
	all	213	280	0.587	0.817	0.751	0.485		
Epoch	GPU_mem	box_loss	obj_loss	cls_loss	Instances	Size	100%	14/14	[01:54:00:00, 8.17s/it]
43/99	0G	0.03868	0.02084	0.01971	14	640	100%	14/14	[01:54:00:00, 8.17s/it]
	Class	Images	Instances	P	R	mAP50	mAP50-95	100%	7/7 [00:33:00:00, 4.75s/it]
	all	213	280	0.587	0.817	0.751	0.485		

Figure 3: Training of the model

Following training completion, the trained batches' results are kept in the final model and annotated with the class numbers. There will be many duplicate detections with overlapping bounding boxes even though we disregarded weak detections. High overlapping box removal is achieved using non-max suppression. Figure 4 shows an example of how patch training display the results for the experiment patches:

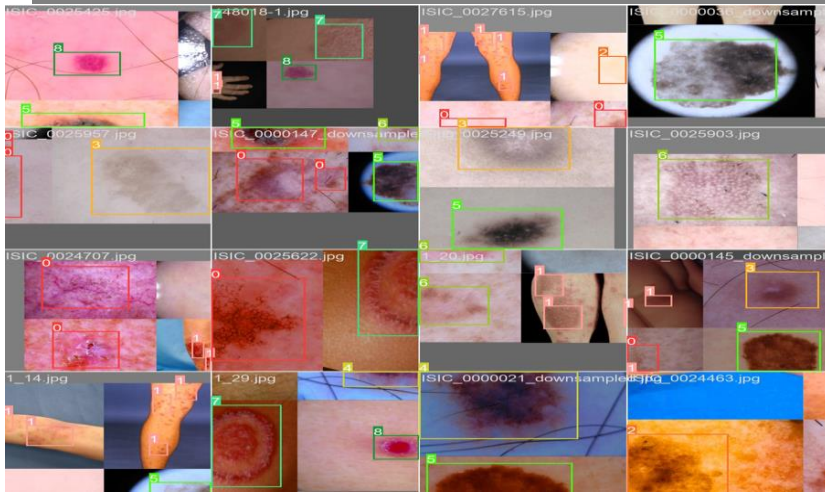


Figure 4: Sample batch training

Figure 5 display the precision and recall with a high accuracy reaches to 99.5% for some diseases:

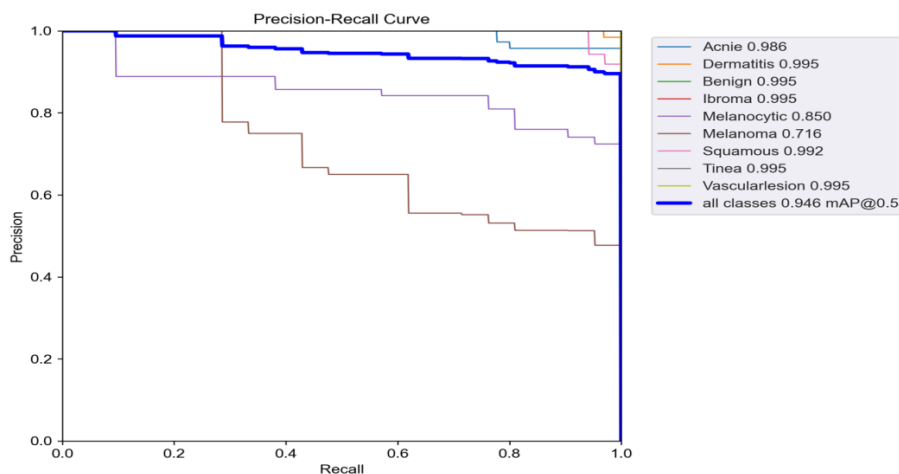


Figure 5: Precision and Recall

Also, in figure 5 the prediction is ranged from 99.5% to 99% for most diseases. Table 1 shows the prediction accuracy that results from the execution of the experiment where the overall score for the model is 98.74%:

Table 1: Model accuracy for the skin diseases – Training results

Disease	Prediction
Acne	98.6%
Dermatitis	99.5%
Benign	99.5%
Ibroma	99.5%
Squamous	99.2%
Tinea	99.5%
Vascularlesion	99.5%
Overall	98.74%

An automated Skin Disease detection system through machine learning using image processing techniques is a rapidly developing field with much research work still under progress.

In the comparison with the previous research which based on model's improvement on sensitivity results using the SVM algorithm (95%) in comparison to the decision tree (93%) and even for the case of the KNN (94%) algorithm, which indicates that the model outperforms on the large dataset using the YOLO [34]. Table 2 shows the results of those previous research which using the HAM10000 datasets:

Table 2: Results of previous used models

Model	SVM	KNN	DT
Accuracy	97	95	95
Precision	97.71	95.71	95.14
Recall	97.57	95.57	95.14
F1-score	97.43	95.14	94.71
Log loss	11.37	15.59	17.37

The skin disease predictions are observed, and the remaining detections are eventually identified by drawing bounding boxes around them and displaying the resultant image. Here is a sample for the prediction of some diseases:

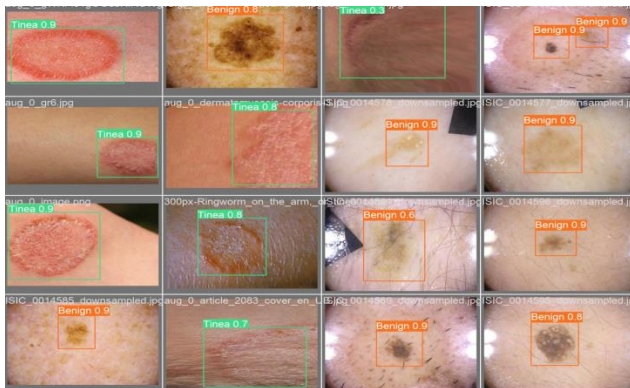


Figure 6: Prediction results after training the proposed model

In addition, the confusion matrix corresponding to the best accuracy of the model training is shown in figure 7:

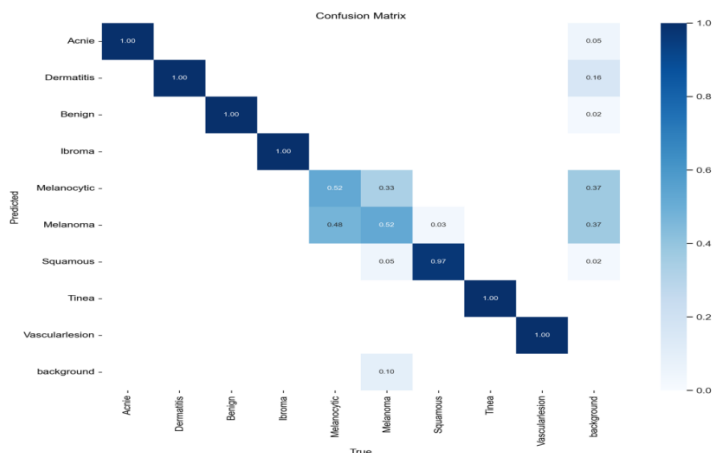


Figure 7: Confusion matrix of the training process

DISCUSSION

Even while our suggested model performed well in classification on datasets with significant imbalances or few samples, it was still far from perfect and contained flaws. For instance, the training and labeling of the photos for the sample preparation in our suggested model required a significant amount of processing power, and the training speed was comparatively slow. As a result, we will be providing more image labeling and more samples in our future work to improve the detection quality, as it has been demonstrated that some diseases are still not correctly diagnosed because of inadequate data sources. Furthermore, we want to evaluate our proposed model on additional benchmark datasets representing various skin conditions.

CONCLUSION

In this study, we created a YOLO model for the classification of skin conditions using DarkNet. This model has been chosen to be the base sub-classification model for our proposed model. In addition, each sub-classification model's core block now includes an attention module, which aids the network in identifying a region of interest and enhances object detection through picture categorization.

The original iteration of YOLO is exceptionally fast, effective, and intuitive. YOLO improves the state of object detection using a new training and deployment framework for PyTorch, even though it does not yet introduce new model architecture upgrades to the YOLO model family. Ultimately, these test findings proved that the classification performance of our suggested model was superior.

REFERENCES

1. Karimkhani, C.; Dellavalle, R.P.; Coffeng, L.E.; Flohr, C.; Hay, R.J.; Langan, S.M.; Nsoesie, E.O.; Ferrari, A.J.; Erskine, H.E.; Silverberg, J.I. (2017). Global skin disease morbidity and mortality: An update from the global burden of disease study. *JAMA Dermatol.* 153, 406–412.
2. Leiter, U.; Eigentler, T.; Garbe, C. (2014). Epidemiology of skin cancer. *Sunlight Vitam. D Ski. Cancer*, 810, 120–140.
3. Baumann, B.C.; MacArthur, K.M.; Brewer, J.D.; Mendenhall, W.M.; Barker, C.A.; Etzkorn, J.R.; Jellinek, N.J.; Scott, J.F.; Gay, H.A.; Baumann, J.C. (2020). Management of primary skin cancer during a pandemic: Multidisciplinary recommendations. *Cancer*, 126, 3900–3906.
4. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88.
5. Abdulrahman, A.A.; Rasheed, M.; Shihab, S. (2020). The Analytic of image processing smoothing spaces using wavelet. In *Proceedings of the Ibn Al-Haitham International Conference for Pure and Applied Sciences (IHICPS)*, Baghdad, Iraq, 9–10 December 2020; p. 022118.
6. Rashid, T.; Mokji, M.M. (2022). Low-Resolution Image Classification of Cracked Concrete Surface Using Decision Tree Technique. In *Control, Instrumentation and Mechatronics: Theory and Practice*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 641–649.
7. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing* 2017, 234, 11–26.
8. Pouyanfar, S.; Sadiq, S.; Yan, Y.; Tian, H.; Tao, Y.; Reyes, M.P.; Shyu, M.-L.; Chen, S.-C.; Iyengar, S.S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Comput. Surv. (CSUR)* 2018, 51, 1–36.
9. Ker, J.; Wang, L.; Rao, J.; Lim, T. (2017). Deep learning applications in medical image analysis. *IEEE Access* 2017, 6, 9375–9389.
10. Anwar, S.M.; Majid, M.; (2018). Qayyum, A.; Awais, M.; Alnowami, M.; Khan, M.K. Medical image analysis using convolutional neural networks: A review. *J. Med. Syst.* 2018, 42, 226.
11. Codella, N.C.; Gutman, D.; Celebi, M.E.; Helba, B.; Marchetti, M.A.; Dusza, S.W.; Kalloo, A.; Liopyris, K.; Mishra, N.; Kittler, H. (2017). 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 *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, DC, USA, 4–7 April 2018; pp. 168–172.

12. Tschandl, P.; Rosendahl, C.; Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. Data* 2018, 5, 180161.
13. Lan, Z.; Cai, S.; He, X.; Wen, X. (2022). FixCaps: An Improved Capsules Network for Diagnosis of Skin Cancer. *IEEE Access* 2022, 10, 76261–76267.
14. Sabour, S.; Frosst, N.; Hinton, G.E. (2017). Dynamic routing between capsules. *Adv. Neural Inf. Process. Syst.* 2017, 30, 3856–3866.
15. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
16. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. (2020). Supplementary material for ‘ECA-Net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 13–19 June 2020; pp. 13-19.
17. Mobiny, A.; Singh, A.; Van Nguyen, H. (2019). Risk-aware machine learning classifier for skin lesion diagnosis. *J. Clin. Med.* 2019, 8, 1241.
18. Wang, S.; Yin, Y.; Wang, D.; Wang, Y.; Jin, Y. (2021). Interpretability-based multimodal convolutional neural networks for skin lesion diagnosis. *IEEE Trans. Cybern.* 2021, 52, 12623–12637.
19. Allugunti, V.R. (2022). A machine learning model for skin disease classification using convolution neural network. *Int. J. Comput. Program. Database Manag.* 2022, 3, 141–147.
20. Anand, V.; Gupta, S.; Koundal, D.; Nayak, S.R.; Nayak, J.; Vimal, S. (2022). Multi-class Skin Disease Classification Using Transfer Learning Model. *Int. J. Artif. Intell. Tools* 2022, 31, 2250029.
21. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
22. Thurnhofer-Hemsi, K.; López-Rubio, E.; Domínguez, E.; Elizondo, D.A. (2021). Skin lesion classification by ensembles of deep convolutional networks and regularly spaced shifting. *IEEE Access* 2021, 9, 112193–112205.
23. Karthik, R.; Vaichole, T.S.; Kulkarni, S.K.; Yadav, O.; Khan, F. (2022). Eff2Net: An efficient channel attention-based convolutional neural network for skin disease classification. *Biomed. Signal Process. Control* 2022, 73, 103406.
24. Hu, J.; Shen, L.; Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
25. Tan, M.; Le, Q. (2021). Efficientnetv2: Smaller models and faster training. In *Proceedings of the International Conference on Machine Learning*, Shenzhen, China, 26 February–1 March 2021; pp. 10096–10106.
26. Abayomi-Alli, O.O.; Damasevicius, R.; Misra, S.; Maskeliunas, R.; Abayomi-Alli, A. (2021). Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold. *Turk. J. Electr. Eng. Comput. Sci.* 2021, 29, 2600–2614.
27. Mendonça, T.; Ferreira, P.M.; Marques, J.S.; Marcal, A.R.; Rozeira, J. (2013). PH 2-A dermoscopic image database for research and benchmarking. In *Proceedings of the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 3–7 July 2013; pp. 5437–5440.
28. Hoang, L.; Lee, S.-H.; Lee, E.-J.; Kwon, K.-R. (2022). Multiclass Skin Lesion Classification Using a Novel Lightweight Deep Learning Framework for Smart Healthcare. *Appl. Sci.* 2022, 12, 2677.
29. Malibari, A.A.; Alzahrani, J.S.; Eltahir, M.M.; Malik, V.; Obayya, M.; Al Duhayyim, M.; Neto, A.V.L.; de Albuquerque, V.H.C. (2022). Optimal deep neural network-driven computer aided diagnosis model for skin cancer. *Comput. Electr. Eng.* 2022, 103, 108318. [Google Scholar]
30. Nawaz, M.; Nazir, T.; Masood, M.; Ali, F.; Khan, M.A.; Tariq, U.; Sahar, N.; Damaševičius, R. (2022). Melanoma segmentation: A framework of improved DenseNet77 and UNET convolutional neural network. *Int. J. Imaging Syst. Technol.* 2022, 32, 2137–2153.
31. Codella, N.; Rotemberg, V.; Tschandl, P.; Celebi, M.E.; Dusza, S.; Gutman, D.; Helba, B.; Kalloo, A.; Liopyris, K.; Marchetti, M. (2018). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv* 2019, arXiv:1902.03368.

-
32. Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection.
 33. Riya Eliza Shaju. (2023). Skin Disease Classification [Image Dataset]. 900 images to classify 9 diseases (80:20 split). <https://www.kaggle.com/datasets/riyaelizashaju/skin-disease-classification-image-dataset/>
 34. Mostafiz Ahammed, Md. Al Mamun, Mohammad Shorif Uddin, "A machine learning approach for skin disease detection and classification using image segmentation", Healthcare Analytics, Volume 2, 2022, 100122, ISSN 2772-4425, <https://doi.org/10.1016/j.health.2022.100122>.