Optimized Object Detection Model for Automated Skin Disease Analysis

Mariam Binte Bashir
Dept. of ECE
North South University
Dhaka, Bangladesh
mariam.binte@northsouth.edu

MD Ilias Bappi
Department of AI Convergence
Chonnam National University
Gwangju, South Korea
i_bappi@jnu.ac.kr

Mithila Arman
Dept. of CSE
BRAC University
Dhaka, Bangladesh
mithila.arman@g.bracu.ac.bd

Kyungbaek Kim
Department of AI Convergence
Chonnam National University
Gwangju, South Korea
kyungbaekkim@jnu.ac.kr

ABSTRACT

This research examines the fully automated detection of skin diseases based on advanced object detection models trained on a customs-annotated dataset consisting of 7100 images, covering conditions like acne, eczema, and psoriasis. Optimized parameters of the model included the use of an auto-tuned optimizer with learning rate annealing and extra hyperparameter fitting, as these contribute to enhanced speed, accuracy, and model generalization during training. In general, the results reveal a very good precision-recall trade-off by the proposed model, although some conditions, for example acne, need further refinement. Results have established the feasibility of an AI driven skin disease detection algorithm for the real-time diagnosis, while future works shall concentrate on the enhancement of the database and the hardware integration process.

KEYWORDS

Skin disease detection, AI, dermatology, object detection, custom dataset, model optimization, healthcare diagnostics

1 INTRODUCTION

Skin is the large organ part of our body that covers and protects our whole body. The skin has many functions like holding fluid, helping to feel sensations, stabilizing body temperature, keep out bacteria, viruses, and other causes of infections. Many types of skin diseases are shown in our body, and between them, common diseases are acne, skin cancer, vitiligo, actinic prurigo, etc. Many reasons are behind this skin disease, such as bacteria, contact with environmental triggers, fungus, viruses and genetics, etc. [1]

AI makes our lives more comfortable. Because people from rural areas did not consider skin diseases, they did not go to the doctor when they saw some rashes. In that case, they put some unknown medicine in that spot. For that reason, fungus spreads on those spots, and sometimes those spots start rotten. Sometimes, in some cases, doctors even fail to identify which type of disease patients have. But AI has brought a huge revolution in this sector, which not only identifies skin diseases but is also able to tell which skin diseases a person has. AI is becoming updated day by day. For that reason, AI can easily predict early skin disease by analyzing anyone's skin. Even by answering some questions, AI can also identify any disease. AI is also able to suggest medicine to people according to their diseases.

2 Literature Review

Skin cancer tops the list of the common cancer issues today, this is per the United State, whereas, the researchers have noticed one out of five American people that have skin cancer during their lifespan suffer from this disease [2-3]. Melanoma is named as the deadliest skin cancer with a mortality rate of 1.62% among other skin cancers [4]. Nevertheless, if identified and treated at the initial phase itself, there happens to be around 14% increase in five-year survival rates [5].

Modern deep learning methods are now the most recognized jobs with superior results in the field of classification, object detection, and segmentation [6-10]. Many scientists proved that now these methods are used not only in simple tasks but also such complex tasks as passing humans [11]. Esteva et al. [12]employed a pretrained convolutional neural network model to enhance the classification system in universal skin disease. They accomplished 0.6 and 0.8 in classification from top-1 and top-3.

On the other side, some researchers are still having an interest in machine learning, CNN, and some old techniques for the skin lesion classification [13-15]. Krizhevsky et al. [16]pointed out that feature extraction is the most important factor that can be used in skin disease detection. Even fewer researchers follow the novel method by integrating machine learning and computer vision with achieving a 95% accuracy rate in six different skin diseases [17]. Some of the researchers are engaged in computer vision whereas some of them suggested a model that reached MioU of 79.46% on the evaluation of a prepared dataset which represents a 15.34% rise when compared to Deeplab v3+ (MIoU of 64.12%) [18]. Besides diagnosing skin disorders, some of the scientists are keen on examining how deeply the disorder has spread across the whole body [19].

Dasari and her colleagues presented a technique for skin diagnostic purposes that serves as the basis of the new framework that covers the detection of the disease [20]. Initially, to segmentation and feature extraction from the infected regions, this framework relies on automatic base classification models. Set up a

PC with an automation program to analyze eczema through the cropping of segments and also to know the degree of the disease. The system comprises three steps, the first one which involves skin detection and subsequent extraction of features like color, texture, borders and, finally, to establish the severity of eczema using Support Vector [21].

3 Our Approach

In our research we use Yolov8 and Yolov9 model for object detection. YOLOv8 also develops more general object representations than Yolov8 and gives a better performance in object detection. Yolov8 can run up to 45 fps without any batch processing on a Titan X GPU while the fast version can run more than 150 fps. This network, YOLOv8, through a single image using all of its parts, forms every possible bounding box. The Jolov8 device has a mechanism that breaks the incoming image into sections of a S × S grid. Yolov8 has succeeded in containing 24 convolutions and two fully connected layers [22]. YOLOv9 has accelerated the process of complex computations while giving high accuracy. Thus, Yolov9 is a faster plus upgraded variant. YOLOv9 gets a 49% exception in the parameters and 43% of exceptions in the computation from YOLOv8 as far as accuracy improvement is concerned, whilst besides that, those improvements also resulted in aiming at a decay of weight parameters plus a cut of power consumption [23]. YOLOv8 and YOLOv9 are applying some innovations like mosaic and mixup augmentation to guarantee the capability of pattern recognition.

4.2 System Design

The optimizer was set to "auto" for training speed and performance improvements in YOLOv8 and YOLOv9 models. In order not to rely on the cosine decay curve for a (smoother) convergence, we also disabled the cosine learning rate (cos_lr=False) to be able to adjust the learning rate more in line with the plots in the previous section. The initial learning rate (lr0) is defined as 0.01 and the final learning rate (lrf) is defined as 0.01 which was calculated as (lr0 * lrf).

We used a momentum of 0.937 to help escape local minima. We also used a weight decay of 0.0005 to reduce overfitting. Dropout was at 0.0 and di different approach to regularization.

Fig. 1, represents our system design and how our models are set and working. This image demonstrates how the Yolo model for object detection should look like. The model has an Input Image that is followed by Preprocessing to ensure the image is ready for analysis. The Stem Convolution outputs basic features and further processes the basic features through three CSP Blocks which capture more complex features through a multi-scale feature extraction process, with transitions bricks that scale the features between the different CSP blocks. The SPPF (Spatial Pyramid Pooling – Fast) layer integrates responsibilities from different scales, which improves the identification of objects of different sizes. Subsequently, FPN Layers fuse features in different levels, and PANet Layers enhancing these features to enhance localization. The Prediction Head then utilizes these features as a way of predicting Bounding Boxes (the location of the objects),

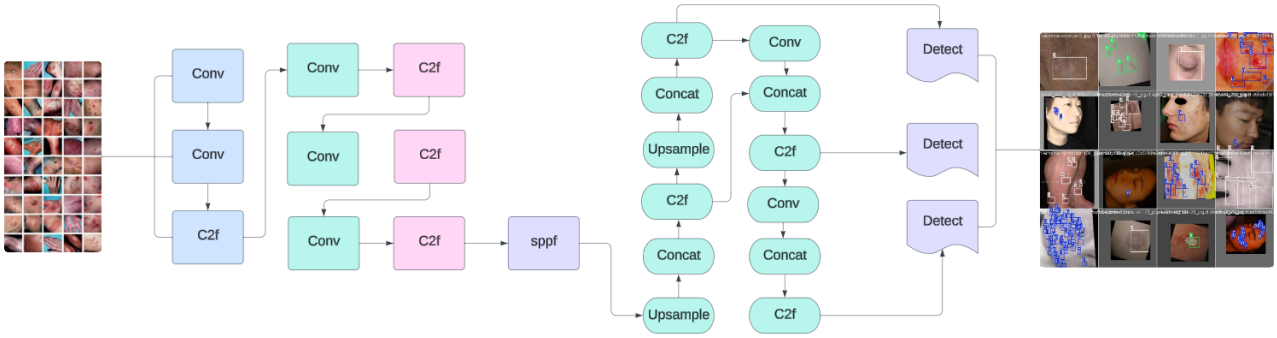

Figure 1: YOLO model architecture.

4 Methodology

4.1 Data Collection

Our research step was the data collection from different datasets and the merging of them into a single dataset. Our total dataset is 7100. Acne, Chicken Skin, Eczema, Hansen-s Disease-Leprosy, Healthy skin, Psoriasis, Ringworm, and warts. All the examples we got are without annotation hence, we ourselves have annotated every single image. However, each image is in 740*740 dimensions.

Confidence scores (the likelihood of objects being actually present in the image), and the Class Output: recognizing objects in the image space.

Fig. 2, We used 3.0 for warmup_epochs for the warmup phase for the learning rate. At this point, a warmup_momentum of 0.8 was introduced during this stage to reduce the speed at which momentum converges to the desired value. Set warmup_bias_lr=0.1 for the bias parameter as the default initial learning rate. Lastly, we set label_smoothing=0.0 to improve generalization while keeping the labels unchanged wrt the ground-truth.

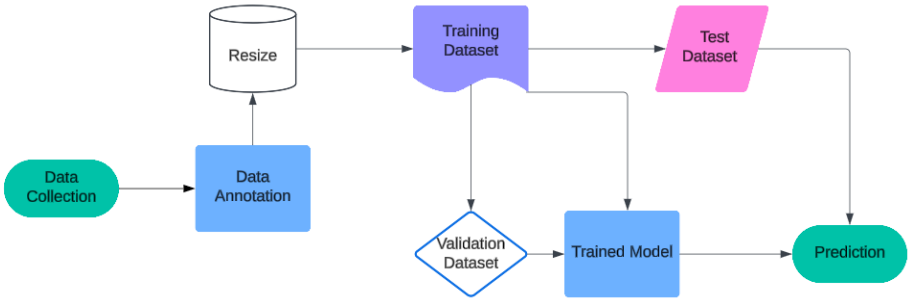

Figure 2: Proposed system framework.

5 RESULTS AND DISCUSSION

Different confidence levels ranging from 0.3 to 0.7 are shown in Figure 2 from a number of annotations alleging "eczema" in various body parts. The skin detection system has shown expected success in locating and describing eczema lesions on skin of different types and on areas. However, the mAP50 mAP50-95 shows that there is still a fair bit to be improved on when it comes to separating out eczema lesions with heightened confidence. The bounding boxes localize the region of interest, suggesting that the model has developed a reasonable understanding of the spatial distribution of eczema. This result indicates that the annotation process succeeded in developing a dataset that allowed the model to generalize among different manifestations of the same skin disease.

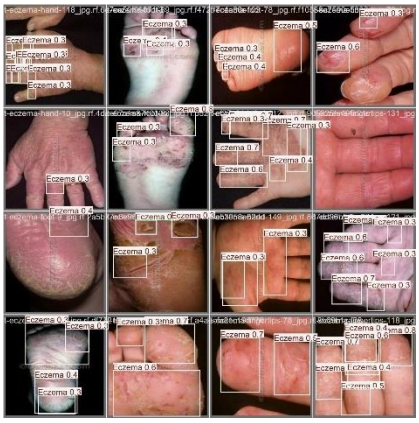

Figure 3: System output from dataset.

Fig. 3, demonstrates the plot of loss against epochs for the bounding box, classification, and DFL (Distribution Focal Loss), as well as precision, recall, mAP50, and mAP50-95 metrics. The plotted curves provide evidence for the model's convergence training. mAP learnt with validation is truly optimistic as such as mAP50 mAP5095, progressively improved, with mAP50 moving towards an asymptotic approach to almost .5 and mAP50-95 simpler than .25. The significances show that the model is competent to localize accurately and classify over a wide range of IoU thresholds, reviving the model strength and encouraging generalization towards novel data. The precision and recall are fluctuating, especially in the early epochs, which is a tendency found in any object detection literature while the model learns to distinguish classes. Box Loss curves for both train and validate datasets experience steady decline, as each epoch records improvement in bounding boxes' prediction by the model. Both training and validation classification loss curves witness a sedimentation in function of time, revealing that the model learned to predict labels for detected objects marvelous.

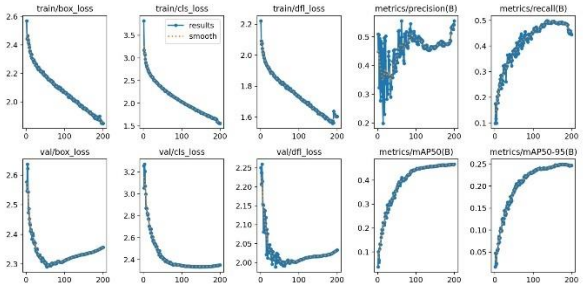

Figure 4: Performance curves across various metric.

This Fig. 4, of acne detection has a confidence score of just 0.29. The low confidence score indicates that the model may have difficulty detecting acne with high confidence possibly because some of the acne lesions share similar visual features with other skin ailments in the dataset.

The Precision-Confidence curve for all classes is illustrated in with eight skin conditions and aggregates performances: Acnes, Chicken Skin, Eczema, Hansen's Disease, Psoriasis, Ringworm, Warts, and Healthy Skin. This curve provides an understanding of the precision-confidence tradeoff by taking a very general global view onto the precision and confidence levels to determine how different conditions are able to trade off these two metrics. A substantially higher degree of confidence displays varying trends in precision between conditions. For instance, both acne and eczema maintain relatively static precision values with higher confidence, while classes like Hansen's disease and warts show more erratic precision as a function of confidence. The blue line reflects the precision at the aggregated level across all classes achieving a precision of 1.0 for a confidence threshold of 0.894, indicating high precision at this confidence level for the existing classes.

Figure 5: Model performance in human body.

These results of Fig. 6, indicate that the YOLO-based model should, prima facie, be capable of detecting many sorts of skin conditions with a balance in precision and recall. The intricacies of the training loss curves attest to effective convergence. The metrics confirm that the model generalizes fairly well onto the test set. However, the low confidence score of certain conditions such as acne and variation along precision-confidence curves representing other classes appropriately indicate the areas needing improvement.

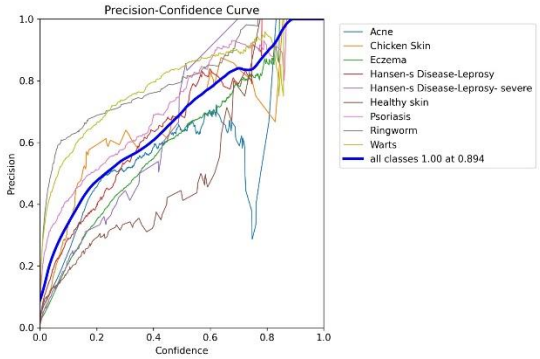

Figure 6: Precision vs. Confidence Curve for Skin Condition Detection across Different Classes.

The enhancement of the dataset, most importantly, with classes diverging in their trend of precision behaviors, should increase confidence in the model's detection for the given samples. In addition, taking on cutting-edge regularization techniques, playing with hyperparameters, or increasing the size and diversity of the dataset may develop more robust results by replenishing the issues that were sparked out.

6 Limitations and Future Work

During our work, we noticed that our model's training accuracy was perfect, but when we tested it in humans, it could have worked better, and it gave low accuracy. Therefore, our models require further fine-tuning to improve accuracy when deployed on hardware in the future. For hardware, we plan to use a camera module and Raspberry Pi5. Because Raspberry Pi5 is perfect for this model implementation, we will also focus on increasing our dataset because we work on nine kinds of skin diseases here. Still, if we look at our dataset, the amount of data is only 7100, which is insufficient, so we will try to increase our dataset by merging more datasets. In the future, we also try to add NLP data to our dataset so that our model not only detects skin disease but also can tell us about the stage and condition of the disease.

7 Conclusion

Nowadays it is must have more attention towards the early detection of different types of skin diseases before they spread to the other parts of the human body. The system we proposed is a skin disease detection system. This system uses images from real time people by capturing picture and identify which type of skin disease that person have. Our experimental result maybe not good at that moment but it is still able to detect skin disease in real time from any human. We will continue modify our system to benefit doctors in real-world clinical practice in dermatological and likely many other types of human disorders.

ACKNOWLEDGMENTS

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2024-RS-2024-00437718, 50%) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation). This work was supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(IITP-2024-RS-2022-00156287, 50%).

REFERENCES

[1] Cleveland Clinic, “Skin Diseases: Types of, Symptoms, Treatment & Prevention,” *Cleveland Clinic*, Jun. 10, 2021. https://my.clevelandclinic.org/health/diseases/21573-skin-diseases

[2] J. Arevalo, A. Cruz-Roa, V. Arias, E. Romero, and F. A. González, “An unsupervised feature learning framework for basal cell carcinoma image analysis,” *Artificial Intelligence in Medicine*, vol. 64, no. 2, pp. 131–145, Jun. 2015, doi: https://doi.org/10.1016/j.artmed.2015.04.004.

[3] S. Pathan, K. G. Prabhu, and P. C. Siddalingaswamy, “Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review,” *Biomedical Signal Processing and Control*, vol. 39, pp. 237–262, Jan. 2018, doi: https://doi.org/10.1016/j.bspc.2017.07.010.

[4] G. Pellacani and S. Seidenari, “Comparison between morphological parameters in pigmented skin lesion images acquired by means of epiluminescence surface microscopy and polarized-light videomicroscopy,” *Clinics in Dermatology*, vol. 20, no. 3, pp. 222–227, May 2002, doi: https://doi.org/10.1016/s0738-081x(02)00231-6.

[5] R. A. Scolyer, S. W. Menzies, C. Longo, Anna Maria Cesinaro, and G. Pellacani, “In Vivo Confocal Microscopy for Diagnosis of Melanoma and Basal Cell Carcinoma Using a Two-Step Method: Analysis of 710 Consecutive Clinically Equivocal Cases,” vol. 132, no. 10, pp. 2386–2394, Oct. 2012, doi: https://doi.org/10.1038/jid.2012.172.

[6] A. Lomas, et al., "A systematic review of worldwide incidence of nonmelanoma skin cancer," *Br. J. Dermatol.*, 2012.

[7] A.-R. A. Ali and T. M. Deserno, "A systematic review of automated melanoma detection in dermatoscopic images and its ground..."

[8] T. P. Habif, et al., *Skin Disease E-Book: Diagnosis and Treatment*, 2017

[9] J. D. Whited, et al., "Does this patient have a mole or a melanoma?" *JAMA*, 1998.

[10] J. L. G. Arroyo and B. G. Zapirain, "Automated detection of melanoma in dermoscopic images," in *Computer Vision Techniques for...*

[11] 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: https://doi.org/10.1016/j.neucom.2021.08.096.

[12] Y. Bengio, et al., "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013.

[13] S. Bhadula, S. Sharma, S. Juyal and C. Kulshrestha, "Machine learning algorithms-based skin disease detection", *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 2, pp. 4044-4049, May 2019.

[14] B. Shetty, R. Fernandes, A.P. Rodrigues, R. Chengoden, S. Bhattacharya and K. Lakshmanna, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network", *Scientific Reports*, vol. 12, no. 1, pp. 18134, Oct 2022.

[15] R. Sumithra, M. Suhil and D. S. Guru, "Segmentation and classification of skin lesions for disease diagnosis", *Procedia Computer Science*, vol. 45, pp. 76-85, Jan 2015.

[16] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", *Advances in Neural Information Processing Systems*, pp. 25, 2012.

[17] J. Premaladha and K. S. Ravichandran, "Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms", *Journal of Medical Systems*, vol. 40, pp. 1-12, April 2016.

[18] . Bekmirzaev, S. Oh, and S. Yo, "RethNet: Object-by-Object Learning for Detecting Facial Skin Problems," Thecvf.com, 2019, Accessed: Nov. 03, 2024. [Online].Available: https://openaccess.thecvf.com/content_ICCVW_2019/html/VRMI/Bekmirzaev_RethNet_Object-by-Object_Learning_for_Detecting_Facial_Skin_Problems_ICCVW_2019_paper.html

[19] O. Trabelsi, L. Tlig, M. Sayadi, and F. Fnaiech, "Skin disease analysis and tracking based on image segmentation," *2013 International Conference on Electrical Engineering and Software Applications*, pp. 1–7, Mar. 2013, doi: https://doi.org/10.1109/iceesa.2013.6578486.

[20] D. A. Reddy, S. Roy, S. Kumar, and R. Tripathi, "A Scheme for Effective Skin Disease Detection using Optimized Region Growing Segmentation and Autoencoder based Classification," *Procedia Computer Science*, vol. 218, pp. 274–282, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.009.

[21] N. Alam, T. Munia, K. Tavakolian, V. Vasefi, N. MacKinnon, and R. Fazel-Rezai, "Automatic detection and severity measurement of eczema using image processing," *IEEE*, 2016.

[22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," May 2016. Available: https://arxiv.org/pdf/1506.02640.

[23] M. Yaseen, "What is YOLOv9: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector," *arXiv.org*, 2024. https://arxiv.org/abs/2409.07813