YOLOv8's head-layer Performance Comparison for Skin Cancer Detection

Deni Sutaji[1,3,*], Oktay Yildiz[2]

[1] *Computer Science Department, Informatics Institute, Gazi Univesity, Ankara, Turkiye*
[2] *Computer Engineering Department, Engineering Faculty, Gazi University, Ankara, Turkiye*
[3] *Informatics Engineering Department, Engineering Faculty, Muhammadiyah Gresik University, Gresik, Indonesia*
[*] *Corresponding author. Email: deni.sutaji@gazi.edu.tr*

ABSTRACT
Skin cancer is a type of cancer that can lead to death. The mortality rate from this disease is high. Detecting the disease at an early stage is essential to prevent the worst impact. However, detection by a dermatologist is time-consuming and costly. Computer Aided Detection (CAD) systems are that aid dermatologists in the understanding of medical images. Deep learning strategies are repeatedly employed in CAD systems. Yolo is one of the well-known deep learning models used to solve detection cases for small, medium, and large objects. In this study, we evaluate the performance of Yolov8 for skin cancer detection considering three head-layers on the HAM10000 dataset. Experimental results show that the head layers with large object paths produce the best mAP and significant speed compared to medium and small. With these results, it can be considered as a reference in future research that using the big object detection path to detect skin cancer at an early stage is recommended.

Keywords: *yolov8, skin cancer, detection, deep learning, head-layer.*

1. INTRODUCTION

The skin has a greater disease risk as the organ with the largest and outermost human body area. One of the causes is a lack of hygiene and intensive exposure to the sun's ultraviolet rays. However, humans frequently miscalculate the diseases on the skin, even skin cancer. In the early symptoms, detecting that the marks or dots on the skin are cancer seeds is challenging. In addition, skin cancer is also one of the diseases that spread very quickly in the body. Moreover, in the worst conditions, skin diseases can cause death for sufferers. So, to minimize this risk, it is necessary to detect cancer early. Dermatologists conduct this detection by visual observation and lab tests. However, this comes with a time-consuming and high cost. Furthermore, the efficacy of detection results is far from satisfactory, exclusively achieving 25% [1].

To overcome these problems, researchers had been developed the skin cancer detection using computer vision and artificial intelligence [2]–[8]. Several studies consider a dermoscopy modality [9], the others utilize a histopathology modality [10], [11] and a few using infrared modality in their proposed methods [12]–[15]. Dermoscopy images are more widely used compared to other modalities [16]–[19].

Convolutional Neural Networks (CNN) and their variations are commonly considered as skin cancer detection approaches [5], [6], [8], [20]. These CNN algorithms generally can be briefly categorized into two specific classes: two-step and one-stage methods [21]. Two-step object detection algorithms generate regions to acquire pre-selected boxes and utilise instance regression and classification with border positioning via CNNs. [21]. Several models were widely considered the two-steps, i.e., R-CNNs[22], Fast-CNNs[23], and Faster R-CNNs [24]. The advantages of these methods are high accuracy performance in the detection object. However, these models produce inefficient time, memory, and process [25]. On the other hand, the one-step models utilize the backbone feature extraction network to directly encounter and classify the target. In the literature, there are several model detections, i.e., Single-Shot multibox Detector (SSD) [27], CenterNet [28], and You-Only-Look-Once (YOLO) [26]. The advantages of these models are low memory dependence and fast results [25]. However, the main limitations are prone to false detection and missing detection.

Yolov8 has three detection paths in the head layer. Each paths have a different pixel size. Therefore, by default, YoloV8 can detect small, medium, and large objects. The object's size is based on the ratio of the object to the image pixel size. P. Hidayatullah et al., in

2022, proposed detecting sperm using only small object paths in detecting sperm movement. It is because the size of the sperm is tiny but has a vast number in an image. The results showed that Yolov5 could work better than a combination of detection paths [29]. Based on their results, in this article we employ the latest version of Yolo to detect skin cancer lesion images. However, as explained earlier, this skin cancer lesion image has more complex challenges, such as: difficult to distinguish lesion and background, various size of cancer lesion, high similarity inter and intra class, and other distracting objects (reeds, bubbles, rulers, paper edges).

In this study, the authors present a comparison performance of Yolov8's head-layers for early skin cancer detection though the following substantial contributions:

(1) Modify Yolov8's head-layers based on object detection paths (small, medium, and big).

(2) Annotation skin lesion from HAM10000 dataset.

(3) Performing transfer learning and fine-tuning by unfreezing the ten last layers of Yolov8's backbone.

The remainder of this paper is outlined as follows: Introduction described in Section 1. Section 2 explained our proposed model and how to evaluate the performance. Results and discussion are comprehensively discussed in the Section 3. To conclude, we summarize our work in section 4.

2. METHODS

In this section we describe our research methodology. In detail, our research steps start from data collection and selection, followed by image label annotation, split data (train, valid, and test), model training, model evaluation, model test, performance analysis. **Figure 1** describes our research methodology.

2.1. Dataset

In the data collection and selection step, we chose the HAM10000 dataset. This dataset has the challenge of class imbalance. Of the 7 skin cancer disease classes, Nevus is the class with the most instances. Thus, we selected 1000 images only. Next, we annotated the HAM10000 dataset images using the labelimg python library. We annotated the image lesions ourselves, based on the cancer class. Based on the lesions we encountered, we had a little difficulty annotating due to the almost equal brightness and luminance of some images between the lesion and the background. In addition, based on our observations, there are similarities in shape, color, and texture of both intra- and extra-class lesions. Furthermore, the size of the lesions in each image also varied. However, as a percentage, there are more lesions with a large object ratio compared to the background. The annotation results are saved using Yolo format and in the form of text files. An example of a lesion image with the annotations we have done is shown in **Figure 2**.

Furthermore, we partitioned the data into training, validation, and test data with a composition of 70:20:10. From a total of 4333 data, we split the training data into 3030, validation data into 862, and test data into 441 images. The details of the data division composition are shown in **Table 1**.

2.2. Yolov8 Modification

As the latest generation of the Yolo family launched by ultralytics in January 2023 [30], version 8 has advantages over previous generations. However, for particular cases, the performance of Yolov8 needs to be modified to get better performance. One such approach is to modify the head layer [29].

There are three detection paths in the head layer of the Yolov8 architecture: small, medium, and large. To determine the performance of each of these paths, we modified the Yolov8 architecture in the head section by separating each path, and then compared the performance of the object detection paths. Each head layer modification is presented in **Figure 3**.

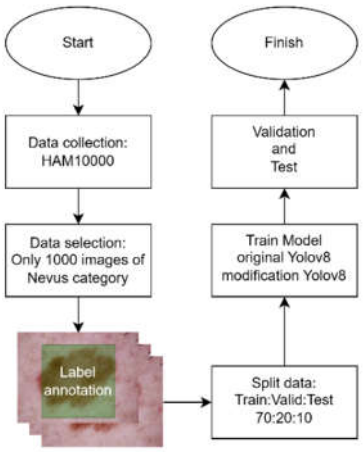

Figure 1. Our proposed methodology

Figure 2. Annotated image manually as melanoma using labelimg python library

Table 1. Dataset distribution (train, valid, and test)

Disease/Class	Total	Train	Val	Test
Akiec	327	228	65	34
Bcc	514	359	102	53
Bkl	1099	769	219	111
Df	119	83	23	13
Mel	1119	783	223	113
Nev	1012	708	202	102
Vasc	143	100	28	15
Total	**4333**	**3030**	**862**	**441**

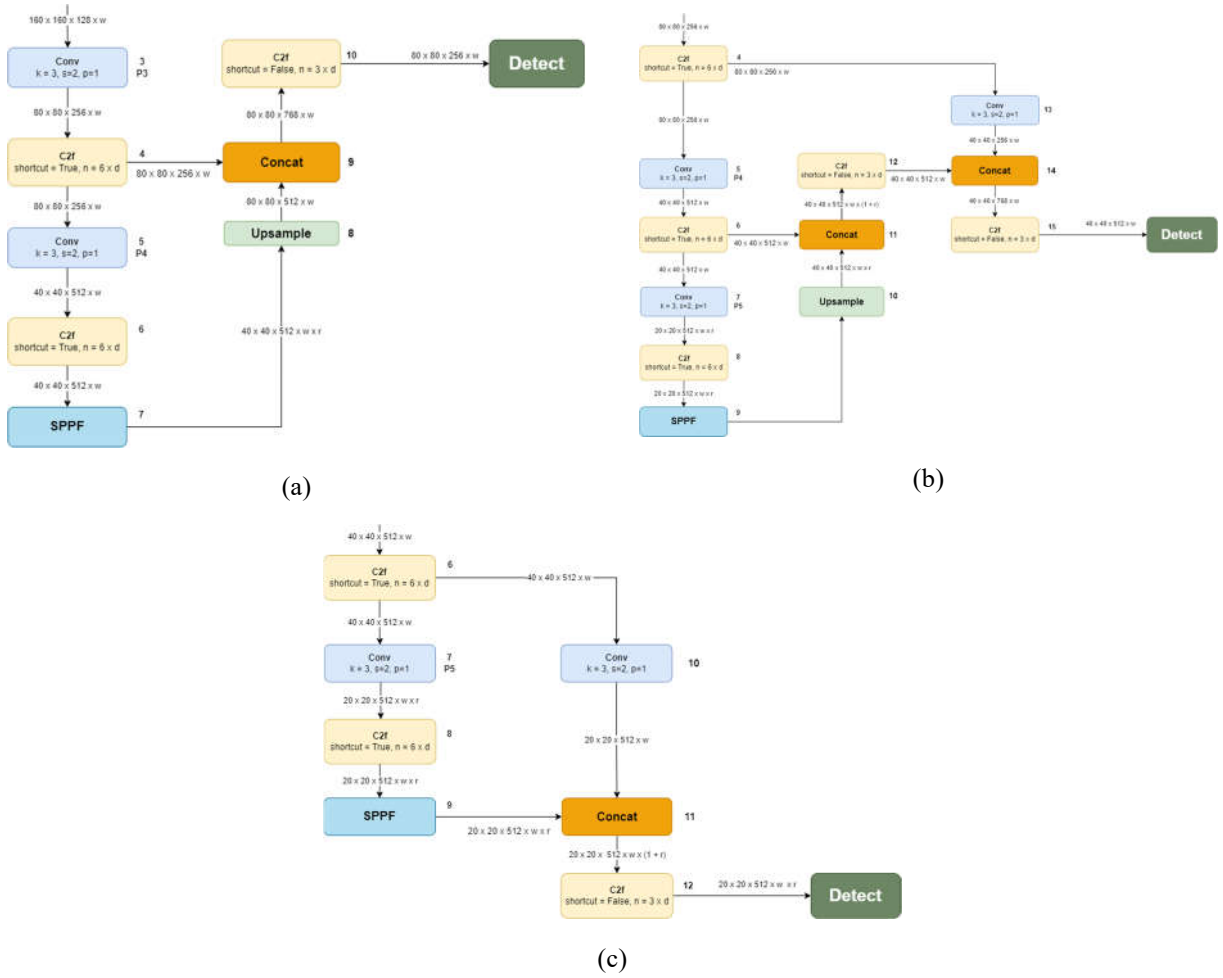

Figure 3. Modification head-layer of YoloV8 (a) Small path object detection, (b) Medium path object detection, and (c) Big path object detection

2.3. Training Yolov8 Modification

Once the data is ready for use, the next step is training the Yolov8 model. We divide the training in 4 stages according to the head-layer evaluation in Yolov8 i.e.: small, medium, big, and combination. During the training process we set the hyperparameters. We set the batch size to 16, epoch 100, pretrained model weights as transfer learning option with the final 10-layer unfrozen.

We used Google Colabs Pro as the service to train the model. The runtime we run has device specifications of V100 GPU with High-RAM speed, 12.7 GB RAM capacity, and multiple CPUs. On the software side, we employ python version 3.10, torch version 2.1 from the Yolov8 installation process using ultralytics library.

2.4. Evaluation indicators

The mean average precision (mAP) is the primary evaluation metric to estimate the detection performance of each path detection in the head-layer modification of Yolov8. This metric is calculated from the intersection over the union (IoU) threshold, explained in **Eq. 1**. IoU can determine correct and incorrect detection. The valid detection condition is retrieved, while IoU of the box

obtained by the model is greater than the threshold. In more detail, mAP is the average of all 10 IoU in the range of 0.50 to 0.95 with an increment value of 0.05 for all specific categories, which is used as the primary metric for ranking. In addition, basic metric Precision and Recall are calculated first before getting the IoU value. These metrics are represented in **Eq. 2** and **Eq. 3**. The mAP is formulated in **Eq. 4**.

$$IoU = \frac{A \cap B}{A \cup B} \quad (1)$$

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

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

$$mAP = \frac{1}{c}\sum_{k=i}^{N} P_{(k)}R_{(k)} \quad (4)$$

3. RESULTS AND DISCUSSION

3.1. Small object path performance

The experimental results for small object path are described in **Table 2**. The model produces the worst performance than others. The small object path was failed detecting the lesion. It caused the ratio of lesion object for dermoscopy images mostly is big and medium size based on human visualization. The best result is obtained from nevus categories.

Table 2. Small path modification performance

Class	Prec	Rec	mAP50	mAP50-95
Akiec	0.094	0.015	0.043	0.011
Bcc	0.665	0.461	0.515	0.179
Bkl	0.361	0.290	0.258	0.112
Df	0.449	0.304	0.304	0.146
Mel	0.398	0.302	0.271	0.126
Nev	0.683	0.587	0.634	0.294
Vasc	0.575	0.357	0.445	0.251
AVG	**0.461**	**0.331**	**0.353**	**0.160**

3.2. Medium object path performance

This modification obtained better results than small path. Detail performance is recorded in **Table 3**. YoloV8 with medium path can detect the lesion with no detection ratio smaller than small path for the test data.

Table 3. Medium path modification performance

Class	Prec	Rec	mAP50	mAP50-95
Akiec	0.411	0.397	0.403	0.241
Bcc	0.532	0.716	0.599	0.321
Bkl	0.411	0.727	0.567	0.374
Df	0.552	0.376	0.408	0.248
Mel	0.482	0.858	0.685	0.523
Nev	0.731	0.817	0.881	0.607
Vasc	0.609	0.500	0.682	0.454
AVG	**0.533**	**0.627**	**0.604**	**0.395**

3.3. Big object path performance

The best result obtained for Yolov8 head-layer modification obtained from the big path object detection. Entire class detection performance illustrated in **Table 4**. The nevus class detection was the best which achieve more mAP50 more than 90%. The worst results were obtained from Df class with mAP50 only 39.4%. Moreover, the ratio of no detection result is significantly improved than small and medium path detection for the training, validating, and testing data, respectively.

Table 4. Big path modification performance

Class	Prec	Rec	mAP50	mAP50-95
Akiec	0.759	0.554	0.694	0.409
Bcc	0.730	0.797	0.830	0.455
Bkl	0.754	0.695	0.780	0.510
Df	0.434	0.522	0.546	0.393
Mel	0.666	0.942	0.897	0.691
Nev	0.945	0.769	0.933	0.639
Vasc	1.000	0.675	0.889	0.565
AVG	0.755	0.708	0.796	0.523

3.4. Performance comparison

Based on the results, the big object path of the head layer of Yolov8 obtained the best performance compared to small and medium. However, the performance of the big object path in mAP50-95 is lighter than the combination of all object paths of Yolov8. Fortunately, training time processing is significantly reduced than path detection in small, medium, and all, respectively. These results prove our hypothesis that we can select the specific object path of YoloV8 to detect the object in skin lesion based on the visual observation that most skin lesions are big ratio in dermoscopy images. This technique provides comparable detection performance to all path detection. The performance of time-processing and mAP are illustrated in **Table 5**.

Table 5. Time-consuming and mAP training comparison

Path	Time (hours)	mAP50	mAP50-95
Small	1.171	0.353	0.160
Medium	1.229	0.604	0.395
Big	**0.598**	0.796	0.523
All	0.828	**0.818**	**0.547**

3.5. Lesion detection performance

This subsection provides an example of skin lesion detection for dermoscopy images. Figure 4 demonstrates the skin lesion detection using big path detection. For example, we present the Bkl, Mel, and Nev classes, which are valid and invalid detection (No detection results).

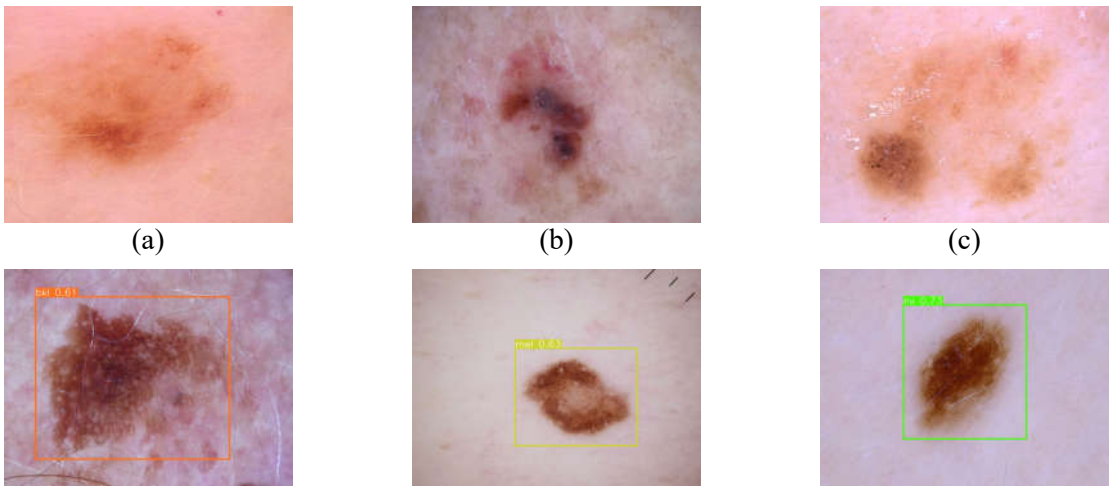

Figure 4. Lesion detection for Bkl, Mel, and Nev, respectively. (a),(b), and (c) No detection results; (d), (e), and (f) Valid detection results

3.4. Limitations

The most challenging in this study is the heterogeneity of skin lesions in size, color, texture, and noise, both intra and extra classes. The best performance of mAP50-95 values less than 55% indicates that the study needs to be improved. As previously mentioned, the annotation process was conducted by ourselves, with minimal knowledge about skin cancer. It might generate the invalid lesion detection processed by the system. However, selecting the path detection in the head layer speeds up the detection process.

For further study, we can involve dermatologists to annotate the skin lesion for improved performance. In addition, we need to modify the Yolov8 architecture from the backbone, neck, and head layer to increase the detection performance [31]. Moreover, ensemble model could be considered to improve the model performance [32].

4. CONCLUSION

Skin cancer detection from the earlier stage utilizing Yolov8 could help the dermatologist diagnose. Experimental results demonstrated that modification Yolov8 in the head-layer considered a big path outperformed that of vanilla Yolov8 and other head-layer path detection, such as small and medium. The small path provided the worst performance, followed by the medium path, and the combination of all paths was second-best. Furthermore, separating the path detection in head-layer significantly speed-up the whole process time, training, validating, and testing, respectively. However, compared with other researches our proposed performance of mAP value is unsatisfactory. It might be caused by poor annotation process.

Further study, we need to refine the annotating skin lesion by assigning a dermatologist to end this issue. Moreover, Other modifications to the yolov8 architecture can be made to the backbone or neck layers for better feature extraction from Yolov8. These can also be done by adding layers, changing the backbone, ensemble two or more architectures, or even reducing layers indicated not to affect performance.

AUTHORS' CONTRIBUTIONS

DS: conceptualization, methodology, investigation, data curation, writing, and editing; OY: supervision, resources, proofreading.

REFERENCES

[1] S. Maurya, S. Tiwari, M. C. Mothukuri, C. M. Tangeda, R. N. S. Nandigam, and D. C. Addagiri, "A review on recent developments in cancer detection using Machine Learning and Deep Learning models," *Biomedical Signal Processing and Control*, vol. 80. Elsevier Ltd, Feb. 01, 2023. doi: 10.1016/j.bspc.2022.104398.

[2] M. K. Hasan, M. A. Ahamad, C. H. Yap, and G. Yang, "A survey, review, and future trends of skin lesion segmentation and classification," *Computers in Biology and Medicine*, vol. 155. Elsevier Ltd, Mar. 01, 2023. doi: 10.1016/j.compbiomed.2023.106624.

[3] Shubhasis Khanra, M. Kuila, S. Patra, R. Saha, and K. G. Dhal, "Survey on

Computational Techniques for Pigmented Skin Lesion Segmentation," *Optical Memory and Neural Networks (Information Optics)*, vol. 31, no. 4, pp. 333–366, Dec. 2022, doi: 10.3103/S1060992X2204004X.

[4] A. R. H. Ali, J. Li, and G. Yang, "Automating the ABCD Rule for Melanoma Detection: A Survey," *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 83333–83346, 2020. doi: 10.1109/ACCESS.2020.2991034.

[5] M. Zafar, M. I. Sharif, M. I. Sharif, S. Kadry, S. A. C. Bukhari, and H. T. Rauf, "Skin Lesion Analysis and Cancer Detection Based on Machine/Deep Learning Techniques: A Comprehensive Survey," *Life*, vol. 13, no. 1. MDPI, Jan. 01, 2023. doi: 10.3390/life13010146.

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

[7] B. Zolfaghari, L. Mirsadeghi, K. Bibak, and K. Kavousi, "Cancer Prognosis and Diagnosis Methods Based on Ensemble Learning," *ACM Comput Surv*, vol. 55, no. 12, pp. 1–34, Dec. 2023, doi: 10.1145/3580218.

[8] J. P. Jeyakumar, A. Jude, A. G. Priya, and J. Hemanth, "A Survey on Computer-Aided Intelligent Methods to Identify and Classify Skin Cancer," *Informatics*, vol. 9, no. 4. MDPI, Dec. 01, 2022. doi: 10.3390/informatics9040099.

[9] U. B. Korkut, Ö. B. Gökturk, and O. Yıldız, "Makine Öğrenmesi ile Bitki Hastalıklarının Tespiti Detection of Plant Disease by Machine Learning," in *26th Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2018.

[10] P. Xie, K. Zuo, Y. Zhang, F. Li, M. Yin, and K. Lu, "Interpretable Classification from Skin Cancer Histology Slides Using Deep Learning: A Retrospective Multicenter Study."

[11] Y. Q. Jiang *et al.*, "Recognizing basal cell carcinoma on smartphone-captured digital histopathology images with a deep neural network," *British Journal of Dermatology*, vol. 182, no. 3, pp. 754–762, Mar. 2020, doi: 10.1111/bjd.18026.

[12] G. Nasreen, K. Haneef, M. Tamoor, and A. Irshad, "Review: a comparative study of state-of-the-art skin image segmentation techniques with CNN," *Multimed Tools Appl*, 2022, doi: 10.1007/s11042-022-13756-5.

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

[14] Y. Wu, B. Chen, A. Zeng, D. Pan, R. Wang, and S. Zhao, "Skin Cancer Classification With Deep Learning: A Systematic Review," *Frontiers in Oncology*, vol. 12. Frontiers Media S.A., Jul. 13, 2022. doi: 10.3389/fonc.2022.893972.

[15] A. Victor, B. S. Gandhi, M. R. Ghalib, and A. M. Jerlin, "A Review on Skin Cancer Detection and Classification using Infrared images," *International Journal of Engineering Trends and Technology*, vol. 70, no. 4. Seventh Sense Research Group, pp. 403–417, Apr. 01, 2022. doi: 10.14445/22315381/IJETT-V70I4P235.

[16] D. Painuli, S. Bhardwaj, and U. köse, "Recent advancement in cancer diagnosis using machine learning and deep learning techniques: A comprehensive review," *Computers in Biology and Medicine*, vol. 146. Elsevier Ltd, Jul. 01, 2022. doi: 10.1016/j.compbiomed.2022.105580.

[17] J. Li, J. Chen, Y. Tang, C. Wang, B. A. Landman, and S. K. Zhou, "Transforming medical imaging with Transformers? A comparative review of key properties, current progresses, and future perspectives," *Med Image Anal*, vol. 85, p.

102762, Apr. 2023, doi: 10.1016/j.media.2023.102762.

[18] A. Mosquera-Zamudio *et al.*, "Deep Learning for Skin Melanocytic Tumors in Whole-Slide Images: A Systematic Review," *Cancers*, vol. 15, no. 1. MDPI, Jan. 01, 2023. doi: 10.3390/cancers15010042.

[19] F. Grignaffini *et al.*, "Machine Learning Approaches for Skin Cancer Classification from Dermoscopic Images: A Systematic Review," *Algorithms*, vol. 15, no. 11. MDPI, Nov. 01, 2022. doi: 10.3390/a15110438.

[20] S. Q. Gilani and O. Marques, "Skin lesion analysis using generative adversarial networks: a review," *Multimed Tools Appl*, Jan. 2023, doi: 10.1007/s11042-022-14267-z.

[21] K. Zhao, R. Lu, S. Wang, X. Yang, Q. Li, and J. Fan, "ST-YOLOA: a Swin-transformer-based YOLO model with an attention mechanism for SAR ship detection under complex background," *Front Neurorobot*, vol. 17, 2023, doi: 10.3389/fnbot.2023.1170163.

[22] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Sep. 2014, pp. 580–587. doi: 10.1109/CVPR.2014.81.

[23] R. Girshick, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Dec. 2015, pp. 1440–1448. doi: 10.1109/ICCV.2015.169.

[24] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." [Online]. Available: https://github.com/

[25] A. Baccouche, B. Garcia-Zapirain, Y. Zheng, and A. S. Elmaghraby, "Early detection and classification of abnormality in prior mammograms using image-to-image translation and YOLO techniques," *Comput Methods Programs Biomed*, vol. 221, Jun. 2022, doi: 10.1016/j.cmpb.2022.106884.

[26] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Jun. 2015, [Online]. Available: http://arxiv.org/abs/1506.02640

[27] W. Liu *et al.*, "SSD: Single Shot MultiBox Detector," Dec. 2015, doi: 10.1007/978-3-319-46448-0_2.

[28] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian, "CenterNet: Keypoint triplets for object detection," in *Proceedings of the IEEE International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers Inc., Oct. 2019, pp. 6568–6577. doi: 10.1109/ICCV.2019.00667.

[29] P. Hidayatullah *et al.*, "DeepSperm: A robust and real-time bull sperm-cell detection in densely populated semen videos," *Comput Methods Programs Biomed*, vol. 209, Sep. 2021, doi: 10.1016/j.cmpb.2021.106302.

[30] J. Terven and D. Cordova-Esparza, "A Comprehensive Review of YOLO: From YOLOv1 and Beyond," Apr. 2023, [Online]. Available: http://arxiv.org/abs/2304.00501

[31] E. Prasetyo, N. Suciati, and C. Fatichah, "Yolov4-tiny with wing convolution layer for detecting fish body part," *Comput Electron Agric*, vol. 198, Jul. 2022, doi: 10.1016/j.compag.2022.107023.

[32] D. Sutaji and O. Yıldız, "LEMOXINET: Lite ensemble MobileNetV2 and Xception models to predict plant disease," *Ecol Inform*, vol. 70, Sep. 2022, doi: 10.1016/j.ecoinf.2022.101698.