Review of Application YOLOv8 in Medical Imaging

Aisyah Widayani [1, a)], Ayub Manggala Putra [1, b)], Agiel Ridlo Maghriebi [1)], Dea Zalfa Cahyla Adi [1)], Moh. Hilmy Faishal Ridho[1)]

[1] *Radiology Imaging Technology, Department of Health, Faculty of Vocational Studies, Universitas Airlangga, Surabaya, Indonesia.*

[a)] Corresponding author: aisyahwidayani@gmail.com
[b)] ayubmanggala.vokasi@drive.unair.ac.id

Abstract. Deep learning has revolutionized medical imaging analysis, with YOLOv8 emerging as a promising tool for various tasks like lesion detection, organ segmentation and disease classification. This review investigates YOLOv8's applications across diverse medical imaging modalities (X-Ray, CT-Scan and MRI). We conducted a systematic literature search across databases like Pubmed, ScienceDirect and IEEE to identify relevant studies evaluating YOLOv8's performance in medical imaging analysis. YOLOv8 achieved high performance for meningioma and pituitary tumors with and without data augmentation (precision >0.92, recall >0.90, mAP >0.93). Glioma detection showed lower performance but still promising results (precision >0.86, recall >0.81, mAP >0.86). Breast cancer detection with SGD optimizer yielded best performance with an average mAP of 0.87 for mass detection. The model achieved high accuracy in detecting normal (mAP 0.939) and malignant lesions (mAP 0.911). YOLO v8 on Dental radiograph successfully detected cavities, impacted teeth, fillings and implants (precision of >0.82, recall of >0.78 and F1-Score of >0.80). Lastly, for lung disease classification, YOLOv8 achieved high accuracy (99.8% training and 90% validation) in classifying normal, COVID-19, influenza and lung cancer disease. With the importance to improve clinical decision-making and patient outcomes in healthcare, the YOLOv8 algorthm underscores the importance of pre-processing, augmentation and optimization of key hyperparameters.

ARTICLE INFO

Article history:

Received 12 March 2024

Revised 18 May 2024

Accepted 31 May 2024

Available online 1 June 2024

Keywords:
Deep Learning,
Object Detection,
YOLOv8,
Medical Imaging

Cite this as:
Widayani, A., Putra, A. M., Maghriebi, A. R., Adi, D. Z. C., & Ridho, M. H. F. Review of Application YOLOv8 in Medical Imaging. Indonesian Applied Physics Letters, 5(1). https://doi.org/10.20473/iapl.v5i1.57001

Indonesian Applied Physics Letters
e-ISSN: 2745-3502
DOI: 10.20473/iapl.v5i1.57001

INTRODUCTION

Computer vision is an aspect of artificial intelligence that can enable computers to perform various tasks such as seeing, feeling, and making decisions. One important part of computer vision is detecting objects. Object detection works by determining the location and position of the object, as well as the label contained in the bounding box (Mahendru and Dubey, 2021). Object detection applications are widely applied in various fields, one of which is the application in the medical field which is used for medical image processing. In medical image processing, identification and localization are needed to find abnormalities in medical images (Qureshi et al., 2023). The assistance of artificial intelligence and deep learning provides opportunities to increase precision and effectiveness in diagnosing various abnormalities in medical imaging.

The deep learning algorithm that is frequently used to detect objects is YOLOv8 (You Only Look Once version 8), with the ability to provide high precision and real-time detection (Palanivel et al., 2023). YOLO is included in the one-stage-detector category, which means it only requires one pass through a neural network and predicts all objects accompanied by bounding boxes (Shetty et al., 2021). The YOLOv8 algorithm is an improved version of the YOLOv3 algorithm, with a wider backbone implementation supported by feature fusion techniques to increase accuracy in detecting objects (Osama, Kumar and Shahid, 2023). The following is the architecture of YOLOv8 in Figure 1.

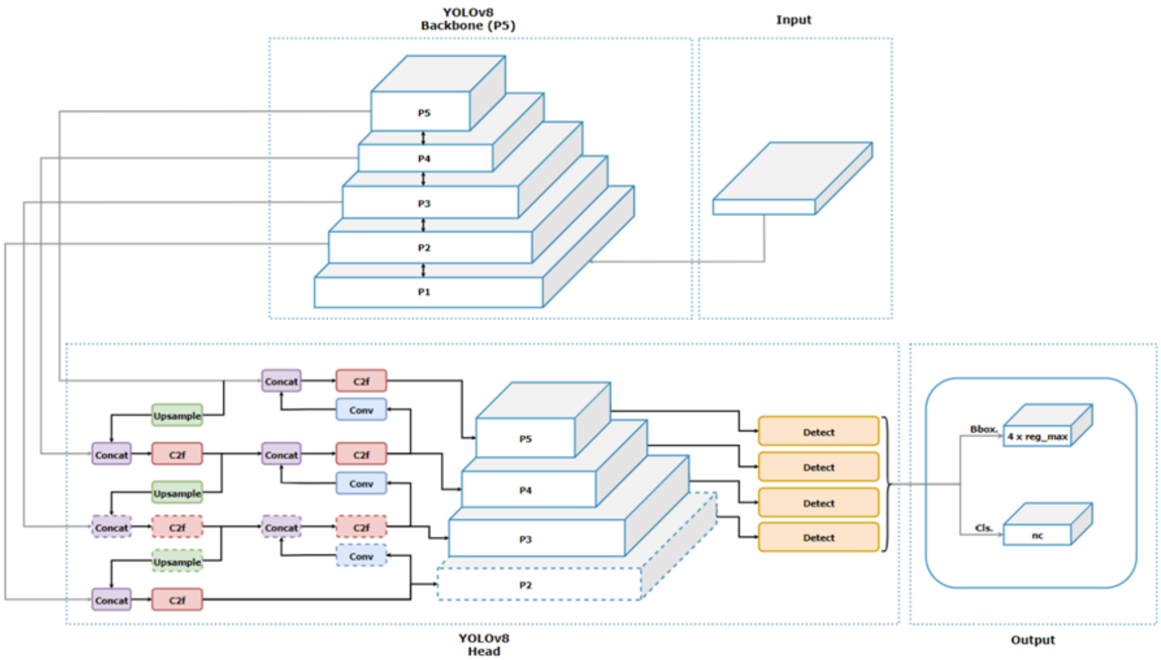


FIGURE 1. YOLOv8 Architecture (Karna *et al.*, 2023)

The YOLOv8 architecture consists of several parts, including CSPDarknet53 Feature Extractor, C2F Module, Detection Head, and YOLOv8-Seg Model (Khare et al., 2023). Each part of YOLOv8 will be discussed one by one.

1. **CSPDarknet53 Feature Extractor** is used as a feature extractor, which consists of convolution layers, batch normalization, and SiLU activation function. Feature extraction was improved in YOLOv8 by changing the convolution layers to 3x3 from the previous 6x6.

2. **C2F Module** implemented in YOLOv8 combines high-level features with contextual information. The combined output of a bottleneck block consists of two 3x3 convolutions with residual connections. The goal is to improve feature representation.

3. **Detection Head** consists of several components, including:

a. **Independent Branches** which allows each branch to carry out certain tasks. This affects the overall detection accuracy.
b. **Activation Function** represents the existence of an object within a bounding box, and allows an object to be classified in a class.
c. **Loss Function** is used to optimize the model using CioU (Complete Intersection over Union) and DFL (Dynamic Focal Loss). Its function is to increase the effectiveness of object detection, especially on small objects.

4. **YOLOv8-Seg Model** is a semantic segmentation model in YOLOv8, so it is called YOLOv8-Seg. This enables broad functionality for a variety of computer vision tasks.

Therefore, this study will discuss several studies that apply the YOLOv8 deep learning algorithm for analysis and diagnosis in medical imaging. The review carried out will support researchers and professionals, especially in the health sector, to understand how the YOLOv8 algorithm works, as well as its benefits in medical object detection applications. In the future, it is hoped that deep learning algorithms for detecting objects can be developed for wider medical applications. This research is composed of Research and Methodology which discusses the approach used and combines data obtained from the included research. The next chapter, related to Results and Discussion, examines the stages carried out before implementing the YOLOv8 algorithm, for example pre-processing and augmentation; then the applied detection class; Data sharing for training, testing and validation processes; implemented hyperparameters, such as epoch, batch, image size, etc; to the training chart results for each study. Then, the last chapter is the conclusion.

RESEARCH METHODOLOGY

This chapter outlines the methodological approach employed to investigate the current landscape of YOLOv8 applications in medical imaging. We conducted a systematic literature search to gather and analyze relevant research studies. This approach ensures a comprehensive and unbiased evaluation of the field's current state and future directions.

Searching Strategy

We employed a multi-database approach, querying relevant publications from PubMed, IEEE and Science Direct. Our search strategy utilized a combination of keywords including "YOLOv8," "medical imaging," "object detection," "segmentation," and "classification" connected by Boolean operators (e.g., "AND," "OR"). We narrowed our search parameters to studies published post-2023 in order to encompass the latest developments in this evolving field, taking into account the release of YOLOv8 in January 2023. After retrieving the initial search results, we implemented a screening process based on pre-defined inclusion and exclusion criteria.

Studies were included if they:
1. Focused on the application of YOLOv8 for medical imaging tasks.
2. Utilized peer-reviewed research methodologies and presented the original research results.
3. Utilized medical imaging datasets acquired from human participants.
4. Evaluated the performance of YOLOv8 using established metrics relevant to the specific medical imaging such as accuracy, precision, recall, F1 score.
5. Released in the English language.

To ensure a targeted analysis of primary research on YOLOv8's application in medical imaging, we implemented the following exclusion criteria:
1. Reviews, editorials, or letters.
2. Studies not published in English.
3. Studies not evaluating YOLOv8.
4. Studies lacking peer-reviewed methodology or original findings.
5. Studies not using established metrics to assess YOLOv8 performance.
6. Studies solely focused on animal models.

Through this process, we identified a collection of relevant studies that formed the foundation for our review and analysis of YOLOv8's contributions to the field of medical imaging.

Evidence Synthesis

Following the initial literature search and application of inclusion/exclusion criteria, the retrieved studies underwent critical appraisal. This involved extracting relevant data points, such as the specific medical imaging modality (e.g., X-ray, CT scan, MRI), the addressed medical imaging task (e.g., lesion detection, organ segmentation), and the reported performance metrics (accuracy, precision, recall).

This extracted data then served as the foundation for a qualitative thematic analysis. This analytical approach aimed to identify recurring themes within the studies, including strengths, weaknesses, and potential research gaps.

RESULT AND DISCUSSION

In this section, we will review several studies using the YOLOv8 model in the field of medical imaging. The first thing that will be discussed is a system diagram which will explain the procedures carried out in the research being evaluated, to find out the steps before implementing the YOLOv8 algorithm. Second, the detection class is the output of the system according to the annotation given to the dataset. Next is the number of datasets used and the division into testing and validation data. Fourth, namely, results related to the accuracy, precision, recall, and F1 Score, which are obtained after the training process as a result of computing. Then, a chart of the training results.

Brain Tumor

The paper entitled "YOLOv8 Based on Data Augmentation for MRI Brain Tumor Detection" (Satila Passa, Nurmaini and Rini, 2023) aims to detect meningiomas, gliomas, and pituitary brain tumors. The data collected from two methods, using without Augmentation and without using Augmentation. The following is a system diagram of the experiments carried out, shown in Figure 2.

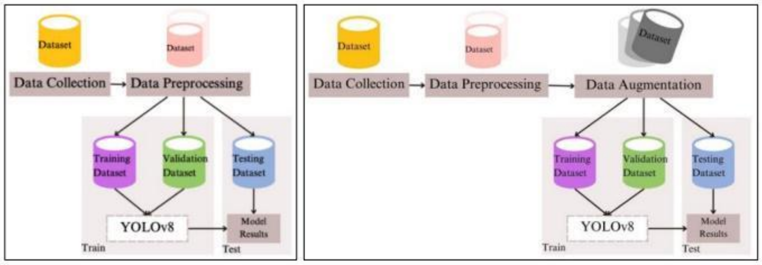


FIGURE 2. (a) Framework without implementing augmentation (b) Implementing augmentation

The first step in this research is collecting data. Data is an important part of creating artificial intelligence (AI) and machine learning (ML). Without relevant and quality data, AI and ML will not be able to produce effective or accurate models. Data taken consist of 3064 T1-weighted contrast-enhanced images with three types of brain tumors such as, meningioma, glioma, and pituitary. The meningioma dataset was separated for training 496, testing 71, and validation 141. Glioma dataset divided into training 998, testing 143, validation 285. Then, 651 training dataset, testing 83, and validation 186 for pituitary.

After getting the required data, the data will be processed. The data processing is to change data from .mat to .jpg format. The next step is data augmentation, which aims to improve the performance of models by training process. There are various techniques applied to this data augmentation such as flip, 90° rotate, crop, rotation, shear, grayscale, brightness, exposure, blur, and noise. Next, training the yolov8 model is performed.

The results of this study evaluate the performance of the YOLOv8 model algorithm for identifying brain tumors, specifically meningioma, glioma, and pituitary tumors. The YOLOv8 hyperparameters used in this research are input size 640 x 640, 100 epochs, and batch size 8. The results obtained in this research are displayed in two without and with data augmentation. The data obtained without and with data augmentation is shown in Figure 3 and 4. Dataset test result that applied YOLOv8 without augmentation shown in Figure 5, with augmentation in Figure 6.

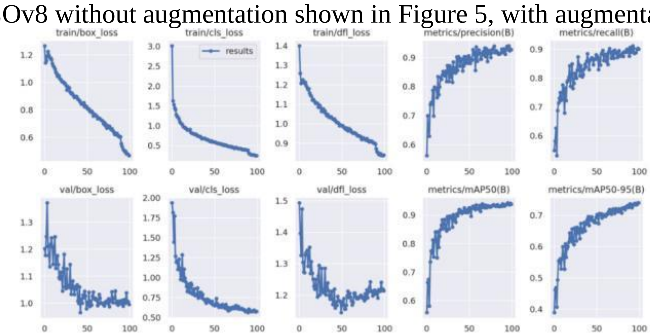


FIGURE 3. Training chart without augmentation

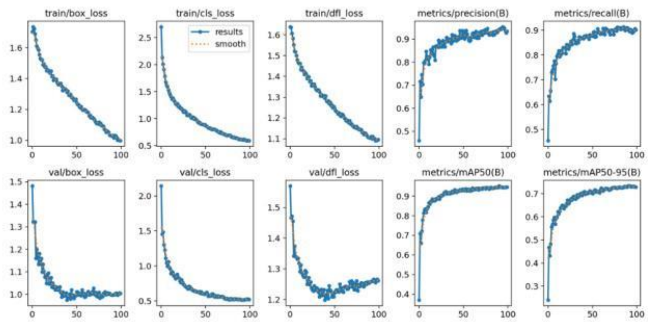


FIGURE 4. Training chart with augmentation

From the data training chart without augmentation (Figure 3), presented for Meningioma tumors, Precision: 0.956, Recall: 0.951, mAP50: 0.98, and mAP50-95: 0.849. Glioma tumors, Precision: 0.866, Recall: 0.816, mAP50: 0.866, and mAP50-95: 0.596. Pituitary tumors, Precision : 0.956, the Recall : 0.939, mAP50 : 0.97, and mAP50-95 : 0.773. Overall performance, Precision : 0.926, Recall : 0.902, mAP50 : 0.938, and mAP50-95 : 0.739.

Figure 4 shows a training chart with augmentation for Meningioma tumors, Precision: 0.985, Recall: 0.95, mAP50: 0.986, and mAP50-95: 0.841. Glioma tumors, Precision: 0.891, Recall: 0.831, mAP50: 0.894, and mAP50-95: 0.599. Pituitary tumors, Precision : 0.95, the Recall : 0.942, mAP50 : 0.975, and mAP50-95 : 0.758. "All" represents overall performance, Precision : 0.942, Recall : 0.908, mAP50 : 0.952, and mAP50-95 : 0.733. Both without and with augmentation, Meningioma dan Pituitary detection get higher precision, recall, and mAP50 scores than Glioma.

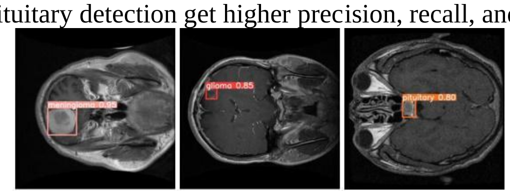


FIGURE 5. Result data testing without augmentation

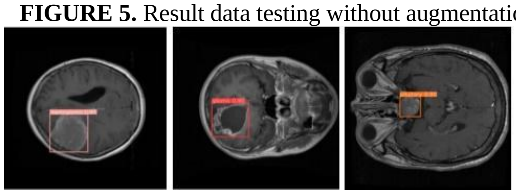


FIGURE 6. Result data testing with augmentation

Breast Cancer

This research (Titisari *et al.*, 2023) focuses on optimizing the YOLOv8s model for breast cancer detection. The input is in the form of a 2D mammography image which is then annotated to indicate the presence of a breast cancer mass or lesion. The next pre-processing stage begins with equalizing the image size by reducing the image pixel size to 640x640. Augmentation is carried out to increase the amount of data so that it is more varied. The respective experiment, the system trained applying three optimization methods, i.e. Adam, SGD, and RMSPropagation restrictions and set up different hyperparameters related to epoch 75, batch size 16. Diagram Research Flow shown on Figure 7.

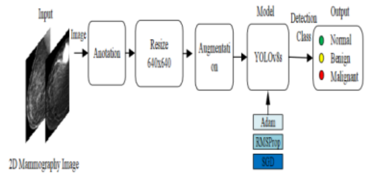
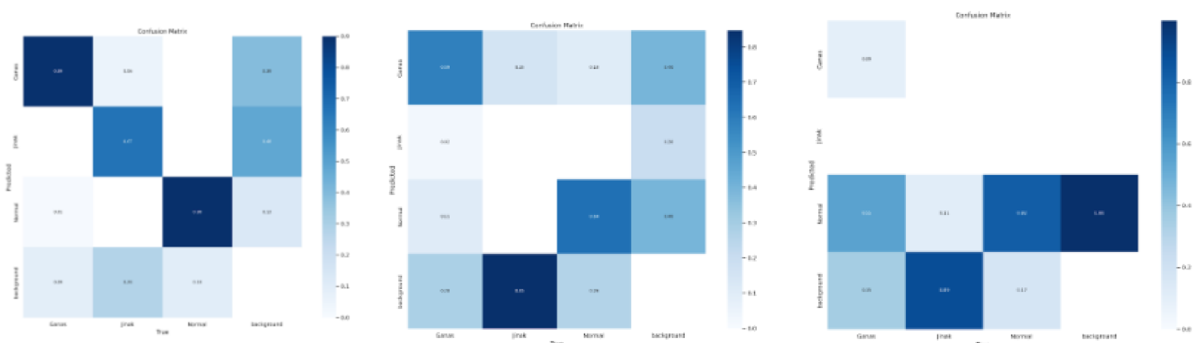


FIGURE 7. Diagram Research Flow of Breast Cancer Detection

The dataset, a mammography image, originating from the Categorized Digital Database for Low Energy and Subtracted Contrast Enhanced Spectral Mammography Images (CDD-CESM), a public dataset. A total of 2,085 mammography images were used after going through a data augmentation process. Of these, there were 809 normal images, 432 benign images, and 844 malignant images. All images are then labeled based on the doctor's medical report on each patient's data. The dataset was then annotated and divided into three subsets, namely 87% as training data, 8% as validation data, and 5% as testing data. In this study, there were a total of 1,824 images used for training, 174 images for testing, and 87 images for evaluation.

The whole work performance of optimizing three different parameters in the YOLOv8s model has been successful in detecting masses or lesions in mammography images. The performance using three optimizers in detecting breast cancer was found that SGD composed **the highest mAP value of 0.87 followed by RMSProp with a value of 0.345 and Adam 0.44.** The SGD optimizer has the fastest training time compared to the other optimizers. Performance of YOLOv8 in detecting normal class, was obtained highest mAP at 0.939 using the SGD optimizer, precision 0.946, and recall of 0.9. The benign class, the highest mAP acquired, utilized the SGD optimizer, precision 0.854, recall 0.667, and mAP 0.762. Whereas, for the Malignant class, the optimum results were also obtained with the SGD optimizer, with a precision of 0.867, recall of 0.871, and mAP of 0.911. Figure 8 shows the confusion matrix of each optimizer.



(a) (b) (c)

FIGURE 8. Confusion Matrix in a) SGD b) ADAM c) RMSProp

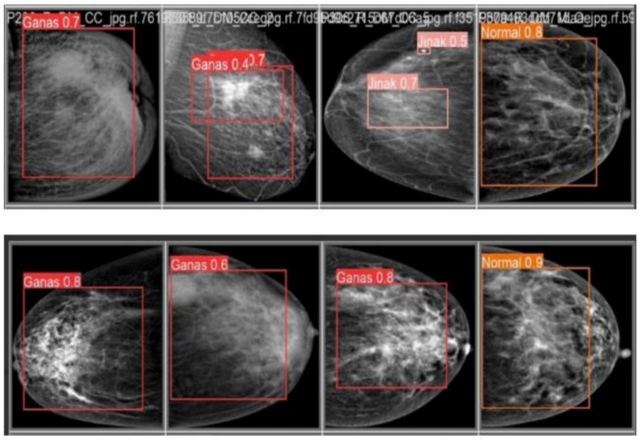


FIGURE 9. Detection Result for Breast Cancer Detection

In Figure 9, There are some classes of detection normal, benign shown as jinak, and malignant displayed as ganas. From the research results, SGD is the best optimizer in detecting masses with an average mAP of 0.87. There was an increase in performance after the tuning process, where accuracy increased from 75% to 85%. The highest mAP value was obtained when detecting the normal (0.939) and malignant (0.911) classes, while when detecting the benign class the value was lower because the amount of benign data was smaller.

Dental Radiography

The application of the YOLOv8 model is also used for analysis and diagnosis in the field of dental radiography. Research (George et al., 2023) states that methods for detecting dental diseases that are carried out manually require time and energy. So, it is important to have an automatic detection system for various dental conditions such as cavities, impacted teeth, fillings and implants on panoramic dental X-rays. The system block diagram from the research carried out is in Figure 10.

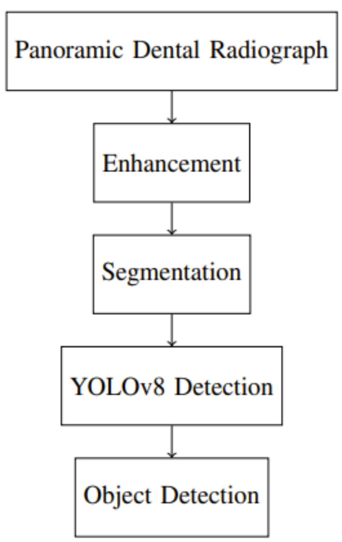


FIGURE 10. Block Diagram of Dental Radiography

Before applying the YOLOv8 model, an enhancement process was carried out to increase the quality of the panoramic dental radiography image by removing noise using a Gaussian Pyramid. In addition, the Laplacian pyramid method is used to increase the detail of the image. Image brightness and contrast are also applied using an automatic algorithm at the enhancement stage.

The segmentation process is carried out to separate four different parts in one image. The segmented parts are divided into extraction of the jaw area, separation of the upper and lower jaw, and vertical segmentation of the upper and lower jaw. The following is an image of an area segmented vertically to separate the upper and lower jaw (Figure .11-14).

FIGURE 11. Top Left Jaw Area

FIGURE 12. Top Right Jaw Area

FIGURE 13. Bottom Left Jaw Area

FIGURE 14. Bottom Right Jaw Area

YOLOv8 is one of the object detection algorithms used to detect the classes determined in this research, namely cavities, impacted teeth, fillings and implants. The number of datasets used is 1269, which are divided into training, testing and validation data, respectively 1015, 63 and 191. There are several hyperparameters used in the training process, namely epoch 35, batch size 16, image size 640, and learning rate 0.001. Precision, recall, and F1-Score are calculated to evaluate the performance of the trained YOLOv8 model. The results obtained for precision were 82.36%, recall 78.38%, and F1 Score 80.32%. The curve results obtained after the training process shown in Figure 15-18, as well as Graphical User Interface (GUI) results which are named "Dental Radiography Analysis and Diagnosis". Users can enter images and the detection process using the YOLOv8 model is shown in Figure 19.

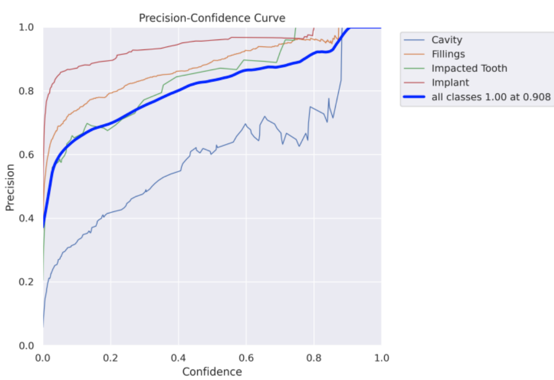


FIGURE 15. Precision-Confidence Curve

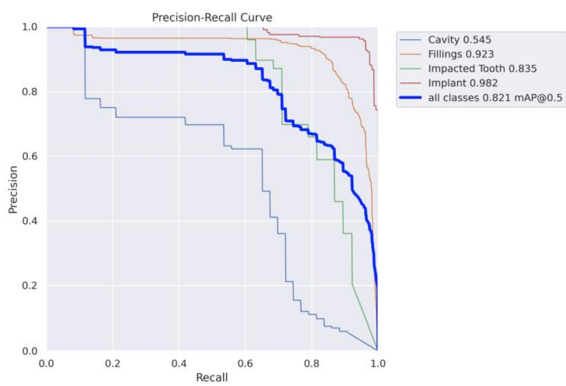


FIGURE 17. Precision-Recall Curve

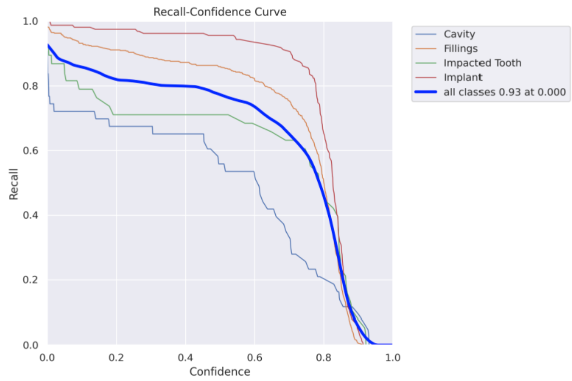


FIGURE 16. Recall-Confidence Curve

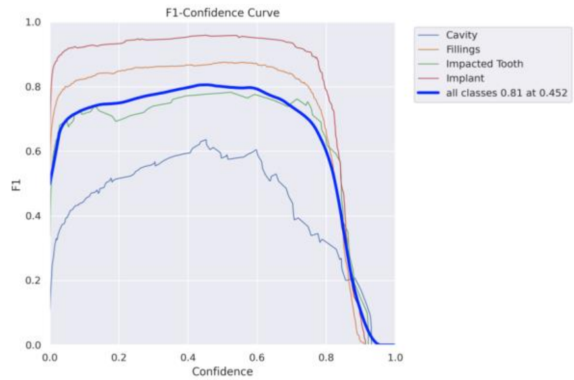


FIGURE 18. F1 Score-Confidence Curve

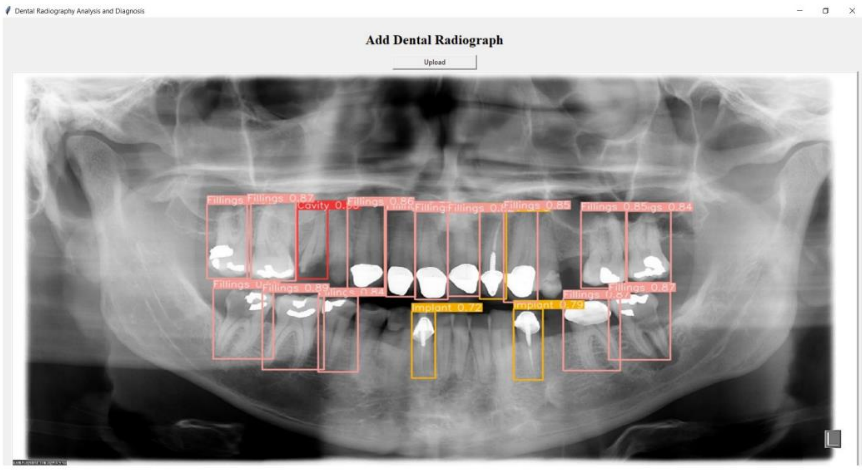


FIGURE 19. GUI of Dental Radiography Analysis and Diagnosis

Lung Disease

In this research, Mousavi *et al.*, 2023 utilized axial lung image results collected from sources such as Kaggle, GitHub, and Radiopedia to import data. The data obtained includes 155 cases of normal, 309 cases of COVID-19, 42 cases of influenza, and 73 cases of cancer. Then, some pre-processing applied to achieve appropriate image quality to avoid ambiguity in detection results. Pre-processing stages are shown in Figure 20.

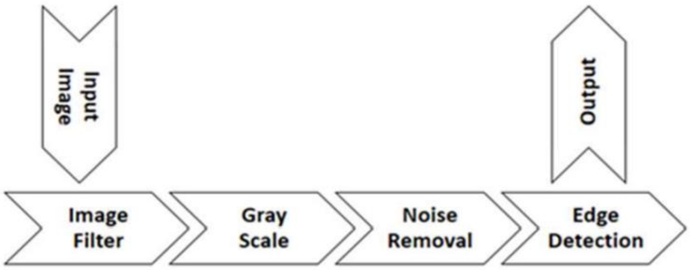


FIGURE 20. Pre-processing stage in Lung Disease Detection

Before the training process, the dataset splits into training, testing, and validation. Normal cases used 100 data train, 50 data testing, and 5 data validation. COVID-19 cases divide to 200 data train, 100 data test, and 9 data validation. Then, Influenza cases separate into 30 data train, 10 data tes, and 2 data validation. Cancer cases distribute data into 50 data train, 20 data tes, and 3 data validation.

There are four types of metrics used to examine the proposed network.i.e. accuracy, precision, recall, and F1-Score. The evaluation metrics included true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) to assess the proposed network. Where, TP represents the number of true positive predictions, while TN denotes the number of true negative predictions. Additionally, FN indicates the number of false negative predictions, and FP indicates the number of false positive predictions. There are such formulas to present accuracy, precision, recall, and F1-Score. That is shown in Equation below.

$$a = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}$$

$$p = \frac{T_P}{T_P + F_P}$$

$$r = \frac{T_P}{T_P + F_N}$$

$$F_{1-Score} = \frac{2 \times precission \times Recall}{precission + Recall}.$$

To determine the accuracy of this study, the researchers used 40 epochs in training. The training accuracy reaches 99.8%, while the validation accuracy achieves 90%. This shows that training provides a very accurate model. Increasing the number of epochs means that it increases the accuracy of the model as well. Loss values recognized are 27% for training, and 0.05% for validation.

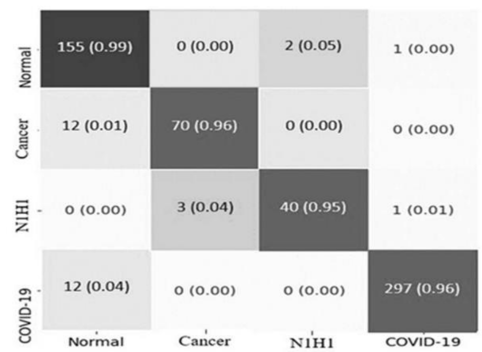


FIGURE 21. Confusion Matrix of Lung Disease Detection

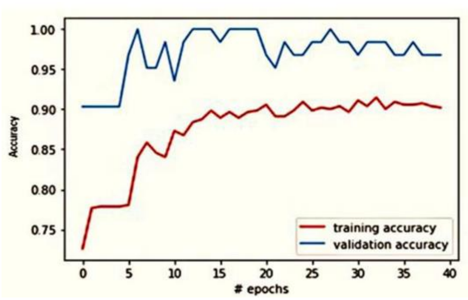


FIGURE 22. The accuracy chart of training models

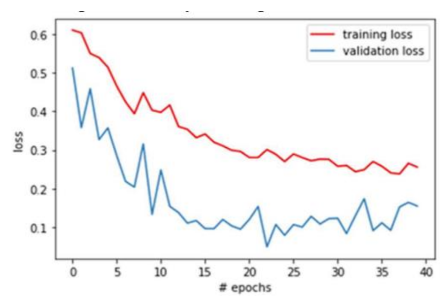


FIGURE 23. The loss chart of training models

CONCLUSION

You Only Look Once version 8 (YOLOv8) algorithm as an object detection model in medical imaging has been applied to detect brain tumors, breast cancer, dental radiography, and lung diseases. Pre-processing has an important role in preparing images with good quality, ensuring that the YOLOv8 algorithm provides accurate information when detecting abnormalities or distinguishing normal conditions. Therefore, this stage is applied in the selected studies. The augmentation process is also significant, aiming to enhance the performance of the YOLOv8 model throughout the training process. The studies discussed herein also delve into hyperparameters, such as epoch, batch size, and image size. Performance metrics including accuracy, precision, recall, and F1-Score are reported to describe the performance results of the YOLOv8 detector, as displayed in the training chart. YOLOv8 demonstrates an important role in enhancing efficiency and accuracy in medical imaging, suggesting its potential to support healthcare in the future. However, continuous research and development are necessary to achieve optimal improvements in healthcare.

REFERENCES

1. George, J. *et al.* (2023) 'Dental Radiography Analysis and Diagnosis using YOLOv8', *9th International Conference on Smart Computing and Communications: Intelligent Technologies and Applications, ICSCC 2023*, pp. 102–107. doi: 10.1109/ICSCC59169.2023.10335023.
2. Karna, N. B. A. *et al.* (2023) 'Toward Accurate Fused Deposition Modeling 3D Printer Fault Detection Using Improved YOLOv8 With Hyperparameter Optimization', *IEEE Access*, 11, pp. 74251–74262. doi: 10.1109/ACCESS.2023.3293056.
3. Khare, O. *et al.* (2023) 'YOLOv8-Based Visual Detection of Road Hazards: Potholes, Sewer Covers, and Manholes', *2023 IEEE Pune Section International Conference, PuneCon 2023*. doi: 10.1109/PuneCon58714.2023.10449999.
4. Mahendru, M. and Dubey, S. K. (2021) 'Real time object detection with audio feedback using Yolo vs. Yolo_V3', *Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering*, pp. 734–740. doi: 10.1109/Confluence51648.2021.9377064.
5. Mousavi, M. *et al.* (2023) 'YOLO for Lung Disease Detection from CT Scans', *SISY 2023 - IEEE 21st International Symposium on Intelligent Systems and Informatics, Proceedings*, pp. 495–500. doi: 10.1109/SISY60376.2023.10417904.
6. Osama, M., Kumar, R. and Shahid, M. (2023) 'Empowering Cardiologists with Deep Learning YOLOv8 Model for Accurate Coronary Artery Stenosis Detection in Angiography Images', *2023 International Conference on IoT, Communication and Automation Technology, ICICAT 2023*, pp. 1–6. doi: 10.1109/ICICAT57735.2023.10263760.
7. Palanivel, N. *et al.* (2023) 'The Art of YOLOv8 Algorithm in Cancer Diagnosis using Medical Imaging', *2023 International Conference on System, Computation, Automation and Networking, ICSCAN 2023*, pp. 1–6. doi: 10.1109/ICSCAN58655.2023.10395046.
8. Qureshi, R. *et al.* (2023) 'A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)', *Authorea Preprints*, 11. Available at: https://www.authorea.com/doi/full/10.36227/techrxiv.23681679.v1?commit=dba07752d065dca931b3a478425ead886b201cf2.
9. Satila Passa, R., Nurmaini, S. and Rini, D. P. (2023) 'YOLOv8 Based on Data Augmentation for MRI Brain Tumor Detection', *Scientific Journal of Informatics*, 10(3), p. 363. doi: 10.15294/sji.v10i3.45361.
10. Shetty, A. K. *et al.* (2021) 'A Review: Object Detection Models', *2021 6th International Conference for Convergence in Technology, I2CT 2021*, pp. 1–8. doi: 10.1109/I2CT51068.2021.9417895.
11. Titisari, D. *et al.* (2023) 'Enhancing Breast Cancer Detection: Optimizing YOLOv8's Performance Through Hyperparameter Tuning', *ICITDA 2023 - Proceedings of the 2023 8th International Conference on Information Technology and Digital Applications*, pp. 1–6. doi: 10.1109/ICITDA60835.2023.10427255.