International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Comparative Study of YOLO Models for Pneumonia Detection

Mohammed Saifuddin Munna[1], Rahul Chowdhury[2], Asif Mohammed Siddiqee[3]

[1,2,3]Department of Electrical & Electronic Engineering, Premier University, Chittagong, Bangladesh

Abstract:
Traditional pneumonia detection methods are usually based on chest X-rays processed with the help of experienced radiologists, allowing for slow processing times and considerable bias. It can have the greatest of implications on the outcome of pneumonia that is a leading killer of children when its timely diagnosis is essential for proper treatment. For accelerated and automated pneumonia identification, deep learning is a prospective technique. Even though different techniques are suggested, some studies proved object detection models to be promising for disguise detection. In this study, we compared three YOLO models (YOLOv3, YOLOv4, and YOLOv6) to determine their performances in detecting pneumonia. We use a dataset of three-class chest X-rays for which we are asked to categorize chest X-rays into normal, viral pneumonia, and bacterial pneumonia. Here, we investigate the performance of different YOLO models to detect pneumonia and classify the type of pneumonia. We aim to demonstrate that the results show the superiority of YOLOv6 over YOLOv3 and YOLOv4, which may help speed up and improve the pneumonia identification in the clinic, ultimately supporting for early intervention and better patient prognosis.

Keywords: Pneumonia, YOLOv3, YOLOv4, Deep Learning, YOLOv6.

1. INTRODUCTION:

Pneumonia, an acute respiratory infection that inflames the air sacs in the lungs, remains a significant global health burden. It claims the lives of an estimated 800,000 children under the age of five annually, making it the leading cause of death in this age group[1]. Early diagnosis and treatment are crucial for improving patient outcomes, particularly in children where the disease can progress rapidly.

Traditional pneumonia diagnosis relies on chest X-rays interpreted by radiologists. While this method provides valuable information, it suffers from limitations. Chest X-ray interpretation can be time-consuming, leading to delays in treatment initiation. Additionally, subjective interpretation by radiologists can introduce variability in diagnosis, potentially impacting patient outcomes.

To address these limitations and improve the efficiency and accuracy of pneumonia detection, researchers have turned to artificial intelligence (AI), specifically deep learning techniques. Deep learning algorithms trained on large medical image datasets have shown promising results in various tasks, including pneumonia detection in chest X-rays[2]. Among these approaches, object detection algorithms have gained significant attention due to their ability to not only identify the presence of pneumonia but also localize the affected regions within the X-ray image [3]

This study focuses on the YOLO (You Only Look Once) family of object detection algorithms, known for their real-time processing capabilities. We aim to conduct a comparative analysis of three YOLO models (YOLOv3, YOLOv4, and YOLOv6) for automated pneumonia detection in chest X-rays. By evaluating their performance in both accurate pneumonia detection and pneumonia type classification, this study seeks to identify the YOLO model that delivers superior results. Our findings have the potential to contribute to the development of a reliable and efficient AI-powered tool for faster and more accurate pneumonia diagnosis in clinical settings, ultimately improving patient outcomes, particularly for children.

2. LITERATURE REVIEW:

Three stages are classified by a supervised learning system in deep learning called pneumonia detection based on YOLO Framework. The COVID-19 epidemic has given rise to more discussion of this subject in recent years[4]. This work aims to detect suspicious objects (Contrast Limited Adaptive Histogram) in chest x-ray photos by applying the You Only Look Once (YOLO) method using CLAHE. Object detection scenarios are commonly combined with YOLO. An S x S grid is created from an image or video before the YOLO algorithm is applied. If an object's center is inside a grid cell, it will detect it. In every grid cell, the bounding box and confidence score are predicted. Yolo computes the probabilities and locations of every class, together with their bounding boxes, simultaneously on every grid. Here are several suggested architectures for a Pneumonia detection system based on the YOLO Framework. The main steps involved in creating our recommended architecture are outlined below, including Make an image dataset that includes a variety of images, The dataset should then be divided into training and testing sets. Annotate the dataset and make any necessary preparations before putting it into the YOLO Framework for the train and test data[5].

Yao et al. propose Pneumonia Yolo (PYolo), an improved YOLOv3 algorithm for pneumonia detection in chest X-rays[6]. PYolo addresses limitations of existing CAD systems by providing lesion location information alongside pneumonia classification. It achieves this through dilated convolution, attention mechanisms, and double K-means clustering, achieving a superior mean average precision (mAP) on a large dataset. This suggests PYolo's potential for more informative and accurate pneumonia diagnosis.

Liu et al. present YOLOv3-P, an improved YOLOv3 for chest X-ray pneumonia detection[7]. Inspired by PANet, YOLOv3-P enhances lesion location via a bottom-up feature fusion pathway. It also utilizes the stronger CSPDarkNet53 backbone for better feature extraction. Achieving a superior mAP (50.43%) on a large dataset, YOLOv3-P offers promise for faster and more accurate pneumonia diagnosis by aiding doctors in lesion localization.

Munna et al. address limitations in chest X-ray interpretation for pneumonia diagnosis. They propose a YOLO-based object detection approach to overcome challenges caused by confounding lung disorders[8]. Their method modifies an existing chest X-ray classification dataset for object detection, enabling localization of pneumonia lesions. The lightweight YOLO model classifies chest X-rays into three categories: Viral Pneumonia, Bacterial Pneumonia, and Normal.

The authors report their method achieves superior performance compared to existing systems, suggesting its potential for faster and more accurate clinical diagnosis, though specific accuracy levels are not mentioned. Kumar et al. propose RYOLO v4-tiny, a lightweight detector for classifying and locating COVID/Non-COVID pneumonia in CT scans and X-rays[9].

This AI model surpasses existing solutions by not only classifying the disease but also pinpointing its location in images, achieving superior performance compared to the baseline YOLO v4-tiny.

Guo et al. [10]proposes MIP-MY, a lightweight YOLOv4 model for tuberculosis detection. MIP-MY reduces model complexity (47% fewer parameters) while improving accuracy (mAP 95.32%) and reducing miss detection (6%) compared to YOLOv4, potentially aiding radiologists in faster and more accurate diagnoses.

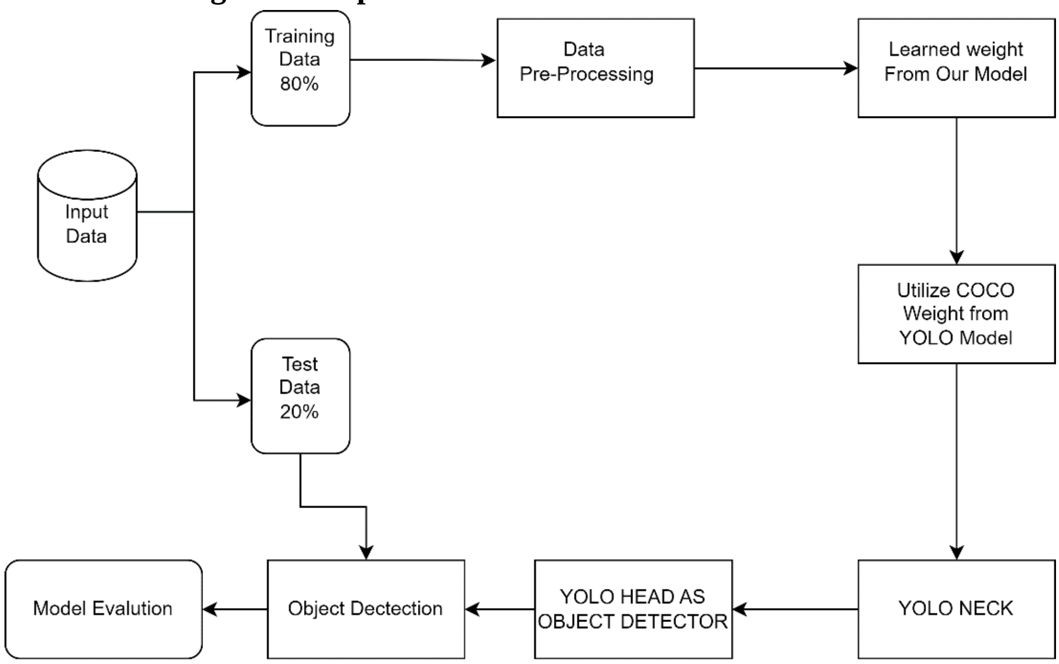
3. METHODOLOGY:

This section provides an overview of the workflow involved in developing the pneumonia detection model. The YOLO framework for pneumonia diagnosis utilizes a supervised machine learning approach. We utilized 2150 photographs from the Kaggle Chest X-ray Images (Pneumonia) dataset, which consists of 5856 X-rays, to annotate data in the YOLO style. The Labelling application software generates a separate text file for every image. The yolo format necessitates a bounding box that precisely delineates the scope of the numerous items linked to a specific class. We utilized three X-ray categories to assign labels to each X-ray.

The dataset was subsequently partitioned into a training set and a test set, maintaining a ratio of 80:20. Subsequently, the training data were divided in the same manner, with 80% allocated for training and 20% for validation. We utilized multiple pre-processing methodologies. The grayscale conversion was implemented to decrease the computational expense.

By employing the normalizing approach, all the pixel values were transformed into a range of 0 to 1. Subsequently, data augmentation techniques, including flipping, rotation, cropping, and scaling, were employed to construct a comprehensive model. Ultimately, the picture standardization method dispersed the pixel values in such a way that they have a mean of 0 and a standard deviation of 1. The block diagram is provided below in Figure 1.

Figure-1 Proposed Model for Pneumonia Detection



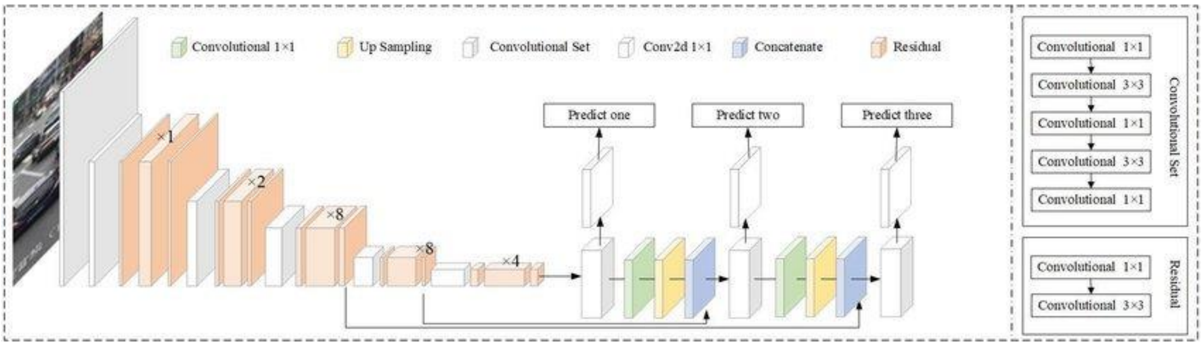
The block diagram shows the steps of a machine learning algorithm for automated pneumonia detection in chest X-rays. It compares the performance of three YOLO models: YOLOv3, YOLOv4, and YOLOv6.

The data is divided into training (80%) and testing (20%) sets. The training data undergoes pre-processing to ensure image quality and consistency. This might involve techniques like resizing, normalization, and data augmentation to improve the model's generalizability.
The pre-processed training data is then used to train the three YOLO models. These models are likely pre-trained on a large image dataset (e.g., ImageNet) and then fine-tuned on the specific chest X-ray dataset for pneumonia detection. During training, the models learn to identify the presence and location of pneumonia regions within the X-ray images. They are also trained to categorize them into normal, viral pneumonia, or bacterial pneumonia based on their visual features.
After training, the models are evaluated on the testing data. Here, the model performance is assessed using standard object detection metrics such as precision, recall, F1-score, and AUC-ROC curve. These metrics evaluate the models' ability to correctly identify pneumonia cases, differentiate between pneumonia types, and avoid false alarms.
The text mentions that researchers aim to identify the YOLO model that delivers superior performance for automated pneumonia detection and classification. This suggests that after evaluating the models on the testing data, the researchers will compare their performance metrics to determine the most accurate model for this task.

3.1 YOLO V3

YOLOv3, also known as You Only Look Once version 3, is a technique for detecting objects in real-time. It achieves this by employing a single convolutional neural network (CNN) to make predictions about the bounding boxes and class probabilities of many objects within an image. YOLOv3, developed by Joseph Redmon and Ali Farhadi in 2018, greatly enhanced the precision and efficiency of object detection in comparison to earlier versions. The model utilizes a Darknet-53 backbone, feature pyramid networks, and a loss function that integrates classification and regression errors. The efficient architecture and outstanding performance of YOLOv3 have led to its widespread adoption in diverse applications such as image identification, video surveillance, and autonomous driving.

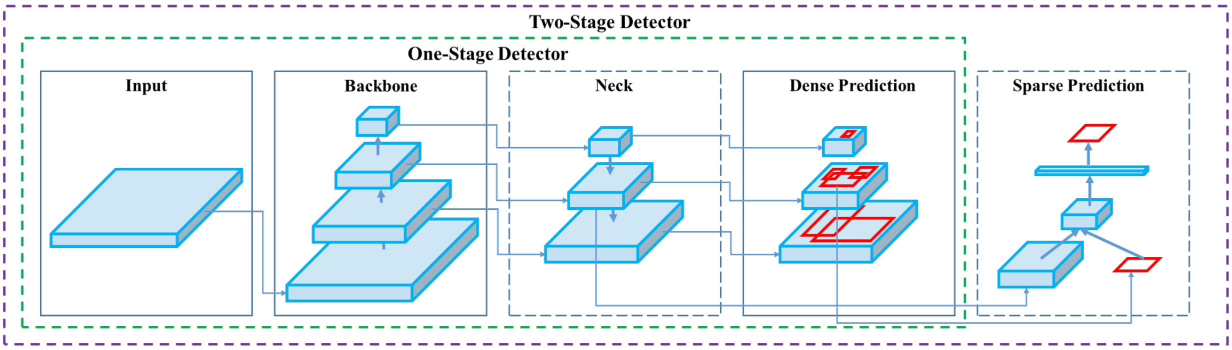
Figure 2. YOLO V3 Model[11]



3.2 YOLO V4

YOLOv4 is a real-time object detection model. It takes an image and performs a single pass through it, unlike some methods that analyze each region multiple times. Internally, YOLOv4 utilizes a pre-trained convolutional neural network (CNN) backbone to extract features, followed by a neck network that merges information from different image resolutions. Finally, a detection head predicts bounding boxes and probabilities for objects within the image. This approach allows YOLOv4 to efficiently identify and localize objects in real-time. An object detector typically consists of numerous components, including the input, the backbone, the neck, and the head. YOLOv4 utilizes a pre-trained model on ImageNet to accurately predict the classes and bounding boxes of objects. The backbone can be derived from many models like as VGG, ResNet, ResNeXt, or DenseNet. The neck component of the detector is responsible for gathering feature maps from several stages. Typically, it consists of many bottom-up pathways and multiple top-down paths. The head component is responsible for doing the ultimate object detections and classifications.

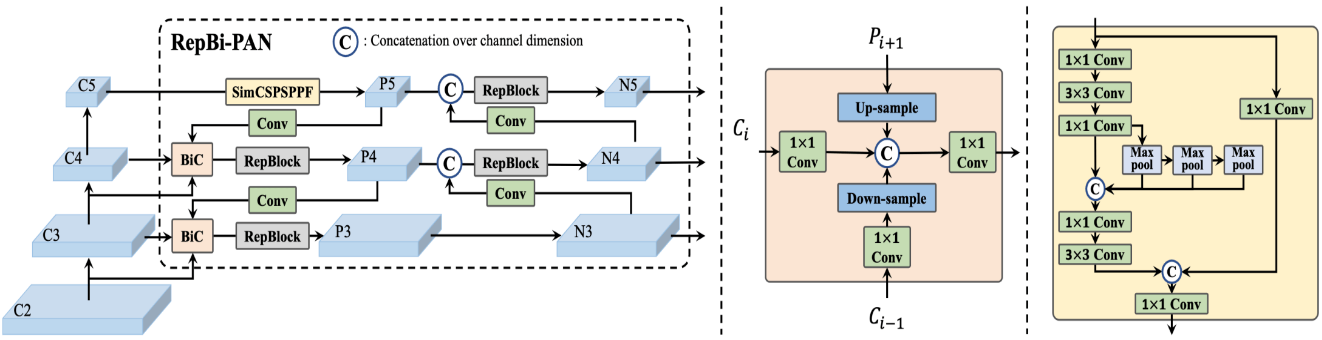
Figure 3: YOLO V4 Model[12]



3.3 YOLOV6

YOLOv6, the newest member of the YOLO family, tackles object detection in a single image pass, aiming for both speed and accuracy. It leverages a pre-trained Convolutional Neural Network (CNN) backbone, often opting for lighter-weight architectures like EfficientNet, to efficiently extract image features. These features are then processed by a specifically designed neck network that focuses on hardware compatibility, allowing for deployment on various platforms. Finally, a decoupled head architecture comes into play. Here, separate prediction heads analyze the features, simultaneously generating bounding boxes for detected objects and assigning probabilities for each object belonging to specific classes. This decoupled approach has the potential to improve detection accuracy. Overall, YOLOv6 prioritizes efficiency for real-time applications while striving for enhanced accuracy through its focus on hardware compatibility and a potentially more powerful prediction architecture.

Figure 4: YOLO V6 Model



(a) RepBi-PAN (b) BiC Module (c) SimCSPSPPF Block

3.4 Data Set:

Publicly available data from the Kaggle Chest X-ray Images (Pneumonia) dataset (https://data.mendeley.com/datasets/rscbjbr9sj/2) was used for this study. The dataset includes 5,856 chest X-ray images of children categorized as viral pneumonia, bacterial pneumonia, or normal. These images have a dimension of 227 x 227 pixels. The dataset is pre-divided into training, validation, and testing sets, with each folder containing images of a specific type: viral pneumonia, bacterial pneumonia, or normal. A sample image from each category is presented in Figure 5. For this research, a subset of 2150 usable images were extracted from the dataset.

Figure 5. Sample Data Set

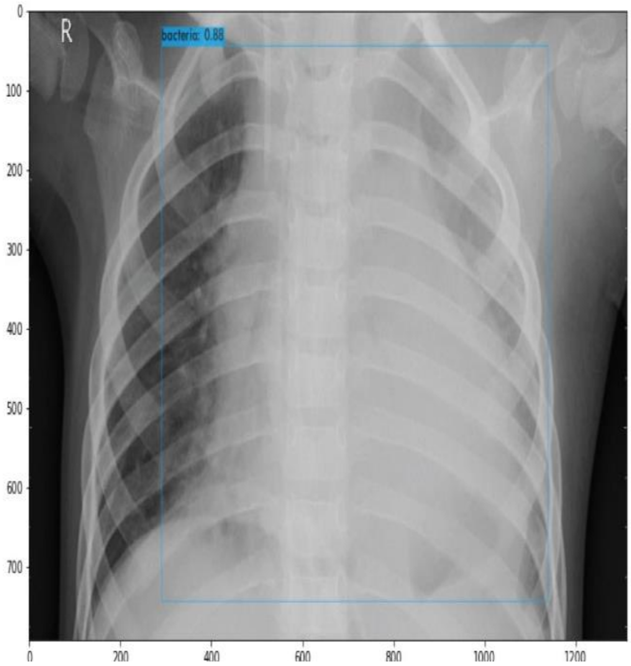
Viral Pneumonia Bacterial Pneumonia Normal Xray

4. RESULT & DISCUSSION:

4.1 Detection of correct Pneumonia Type

Our proposed model yielded two key results. First, it successfully detected pneumonia regions by generating bounding boxes around them. Second, it facilitated a comparative analysis of the accuracy of three models. This analysis revealed an interesting finding regarding processing time. YOLOv3, while effective in detection, exhibited a significantly slower processing speed compared to YOLOv6. This suggests that YOLOv6 might be a more suitable choice for real-time applications due to its faster inference speed. in Figure 6,7,8 detection with boundary box is shown.

Figure 6. Bacterial Pneumonia Detection



International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Figure 7. Viral Pneumonia Detection

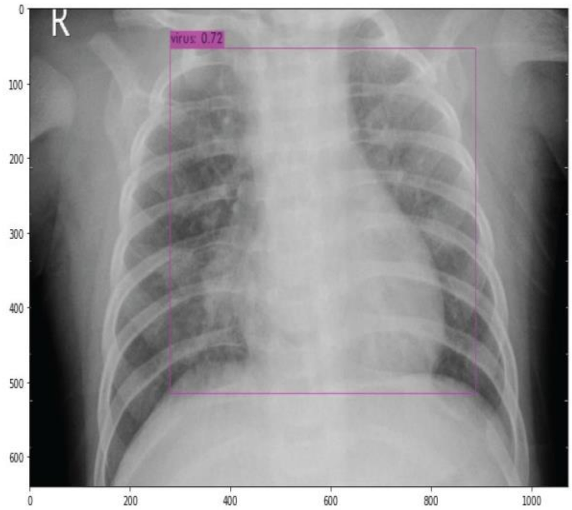
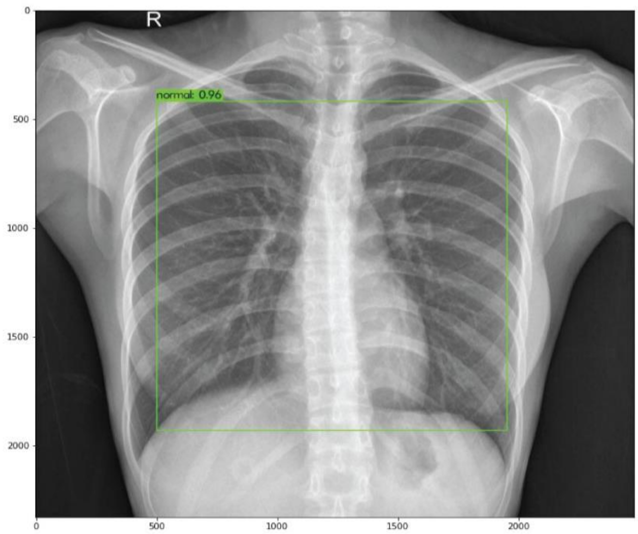


Figure 8. Normal Chest X-Ray Detection



4.2 Comparison of 3 Types of YOLO Algorithm

All three YOLO models (YOLOv3, YOLOv4, and YOLOv6) were trained for a maximum of 6,000 iterations. The resulting performance metrics for each model are presented in Table 1.

Table 1 Comparison of 3 Types of YOLO Algorithm

Model Type	Precision	Recall	F1-Score	Average IoU	MAP
YOLO V3	0.84	0.91	0.90	67.56	93.86
YOLO V4	0.93	0.94	0.94	82.30	96.56
YOLO V6-L	0.95	0.99	0.98	85.19	97.86
YOLO V6-M	0.95	0.99	0.99	88.20	98.86

Figure 9 shows the precision, recall, and F1-score for four different models. The models are YOLOv3, YOLOv4, YOLOv6-L, and YOLOv6-M. The precision is the proportion of true positives among all predicted positives. The recall is the proportion of true positives among all actual positives. The F1-score is the harmonic mean of precision and recall.
The results show that YOLOv6-M has the highest precision, recall, and F1-score among the four models. YOLOv3 has the lowest precision, recall, and F1-score.

Figure 9: Graph of Results (Precision, Recall & F1-Score) for different Models

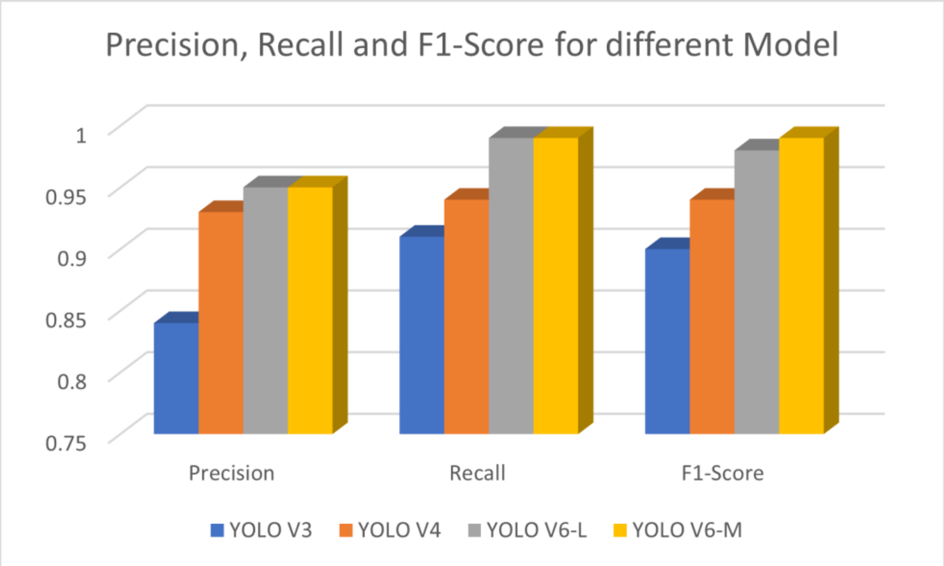


Figure 10: Graph of Results (Average IoU & Mean Average Precision) for different Models

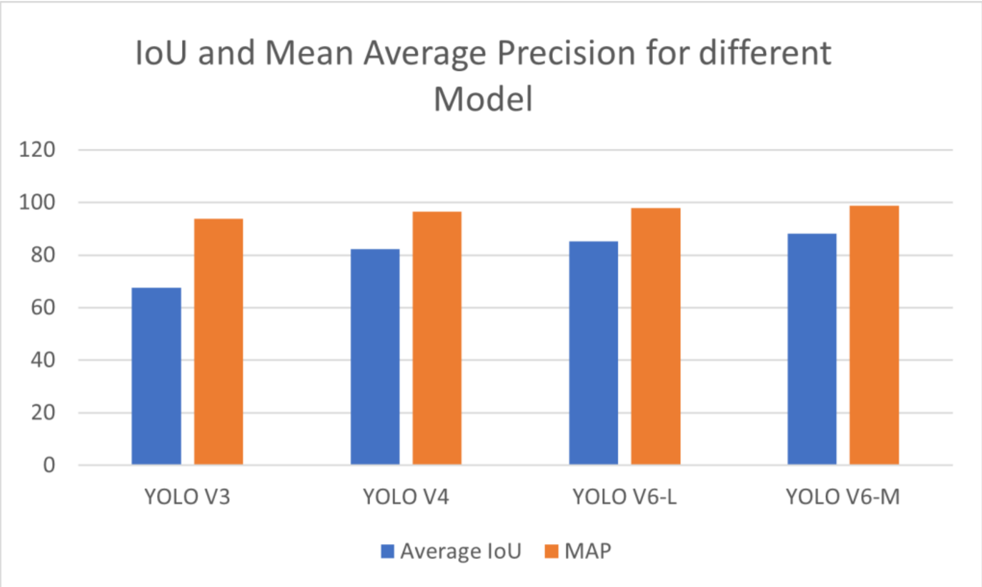


Figure 10 shows the mean average precision (mAP) and intersection over union (IoU) for four different models. The models are YOLOv3, YOLOv4, YOLOv6-L, and YOLOv6-M. The mAP is a measure of how well the model can predict the bounding boxes of objects in an image. The IoU is a measure of how well the predicted bounding boxes overlap with the ground truth bounding boxes. The results show that

YOLOv6-M has the highest mAP, followed by YOLOv3, YOLOv6-L, and YOLOv4. YOLOv6-M also has the highest IoU, followed by YOLOv3, YOLOv4, and YOLOv6-L.
This indicates that YOLOv6-M is the best model for object detection, as it has the highest mAP and IoU.

5. CONCLUSION

A study was undertaken to evaluate the efficacy of YOLO models (perhaps YOLOv3, YOLOv4, and YOLOv6) in automatically detecting pneumonia in chest X-rays. The training process for all three models was limited to a maximum of 6,000 iterations. Afterwards, their performance was assessed using metrics such as precision, recall, and F1-score (shown in Table 1). The examination of these parameters unveiled the important discovery regarding the most efficient approach for pneumonia identification. Furthermore, it was shown that YOLOv6 exhibited a notable superiority in processing speed when compared to YOLOv3.
The results indicate that YOLO models, namely YOLOv6 because of its effectiveness, could be useful instruments in aiding healthcare practitioners to diagnose pneumonia more quickly and possibly with greater accuracy. Additional study is advised to investigate the applicability of these models on larger and more varied datasets. Furthermore, it may be worth considering the integration of domain-specific knowledge into the model design or training process as a means of potentially enhancing detection accuracy.

6. REFERENCES

1. T. Wardlaw, P. Salama, E. W. Johansson, and E. Mason, “Pneumonia: the leading killer of children,” *Lancet*, vol. 368, no. 9541, pp. 1048–1050, Sep. 2006, doi: 10.1016/S0140-6736(06)69334-3.
2. D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, and A. Mittal, “Pneumonia Detection Using CNN based Feature Extraction,” *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*, Feb. 2019, doi: 10.1109/ICECCT.2019.8869364.
3. M. S. Munna and Q. Delwar Hossain, “An Automatic Detection of Pneumonia from Chest Ionizing Radiation Images Using Machine Learning Algorithm,” *2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI)*, pp. 1–5, Dec. 2022, doi: 10.1109/STI56238.2022.10103279.
4. S. Kanakaprabha and D. Radha, “Analysis of COVID-19 and Pneumonia Detection in Chest X-Ray Images using Deep Learning,” *ICCISc 2021 - 2021 International Conference on Communication, Control and Information Sciences, Proceedings*, Jun. 2021, doi: 10.1109/ICCISC52257.2021.9484888.
5. A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” Apr. 2020, Accessed: Jun. 03, 2023. [Online]. Available: https://arxiv.org/abs/2004.10934v1
6. S. Yao, Y. Chen, X. Tian, R. Jiang, and S. Ma, “An Improved Algorithm for Detecting Pneumonia Based on YOLOv3,” *Applied Sciences 2020, Vol. 10, Page 1818*, vol. 10, no. 5, p. 1818, Mar. 2020, doi: 10.3390/APP10051818.
7. H. Liu, J. Cui, and C. Peng, “Pneumonia Detection Algorithm Based on Improved YOLOv3,” *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, vol. 415 LNICST, pp. 313–320, 2022, doi: 10.1007/978-3-030-94182-6_22.

8. M. S. Munna and Q. D. Hossain, “A Lightweight Object Detection Model to Detect Pneumonia Types,” pp. 599–609, 2023, doi: 10.1007/978-981-99-2322-9_45.
9. A. Kumar, “RYOLO v4-tiny: A deep learning based detector for detection of COVID and Non-COVID Pneumonia in CT scans and X-RAY images,” *Optik (Stuttg)*, vol. 268, p. 169786, Oct. 2022, doi: 10.1016/J.IJLEO.2022.169786.
10. Z. Guo, J. Wang, J. Wang, and J. Yuan, “Lightweight YOLOv4 with Multiple Receptive Fields for Detection of Pulmonary Tuberculosis,” *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/9465646.
11. “YOLOv3: Real-Time Object Detection Algorithm (Guide) - viso.ai.” Accessed: Jun. 05, 2024. [Online]. Available: https://viso.ai/deep-learning/yolov3-overview/
12. “YOLOv4 - Ultralytics YOLO Docs.” Accessed: Jun. 05, 2024. [Online]. Available: https://docs.ultralytics.com/models/yolov4/