A Survey on Pneumonia Detection using Computer Vision

Divesh Jadhwani , DYPIU

November 2023

1 Abstract

This comprehensive survey delves into the realm of pneumonia detection through image segmentation, leveraging the YOLO (You Only Look Once) and other CNN based architectures. Pneumonia, a critical respiratory condition, necessitates swift and precise diagnosis. Various deep learning-based segmentation models exist, but this paper posits the advantages offered by YOLO concerning object detection, speed, and adaptability in the context of pneumonia segmentation. The YOLO network is meticulously trained, with pneumonia masks initially converted into polygons for bounding box generation. Subsequently, the model is applied to a validation dataset, generating predicted bounding boxes that are converted back into binary masks via polygon approximation. This paper covers a wide spectrum of methodologies and models, including UNet, Mask R-CNN, ResNet, and ensemble methods, providing insights into their applications for pneumonia detection. The conclusion underscores the significance of computing resources and data quality in enhancing detection, while also highlighting YOLO's remarkable performance with limited data and computational resources.

Keywords: Pneumonia, Image Segmentation, YOLO, Deep Learning, UNet, Mask R-CNN, ResNet, Ensemble Methods, Computer Vision, Data Quality, Computing Resources.

2 Problem Statement

"Can the application of computer vision techniques in medical and healthcare domains reduce the workload of radiologists and enhance the accuracy and efficiency of disease segmentation, particularly in pneumonia diagnosis?"

3 Introduction

Medical image analysis plays a vital role in the early detection and accurate diagnosis of various health conditions. In the context of pneumonia, a respiratory infection that can range from mild to life-threatening, timely and precise diagnosis is crucial. One of the essential tasks in this process is the segmentation of pneumonia regions within medical images, such as X-rays or CT scans. Deep learning-based segmentation models have emerged as powerful tools for automating this critical task. Models like U-Net, ResNet, and others have shown impressive results in diverse medical imaging applications. These models are designed to learn patterns and features within images, making them well-suited for tasks like pneumonia segmentation. However, in this paper, we introduce the YOLO (You Only Look Once) architecture as a noteworthy contender in the field of pneumonia segmentation. While U-Net, ResNet, and similar models are highly effective, YOLO offers distinct advantages.

YOLO's original purpose was object detection, and it excels in this domain. Object detection and image segmentation are closely related because they both involve identifying and precisely locating objects within an image. YOLO's strengths in object detection make it an attractive choice to enhance the precision of pneumonia segmentation. Efficiency is another key feature of YOLO. It provides rapid inference times, which are particularly valuable in the medical domain. In healthcare, quick and accurate diagnoses are critical, and any delay can have serious consequences. Leveraging YOLO allows us to achieve real-time or near-real-time segmentation results. This capability empowers healthcare professionals to make swift and well-informed decisions.[1] In our survey, we present an approaches that utilizes the YOLO architecture and other CNN-based models for pneumonia segmentation. Our approach involves converting pneumonia masks into polygons to train the YOLO model. Subsequently, we apply YOLO to the validation dataset, generating predicted bounding boxes. These predictions are then converted back into binary masks using polygon approximation. This method proves effective, as evidenced by a Dice Score of 0.9479 on the test dataset, underscoring its potential to improve the accuracy and efficiency of pneumonia segmentation in medical imaging.

In summary, computer vision and image segmentation are vital tools in the battle against diseases like pneumonia. While U-Net, ResNet, and other models are well-established choices, YOLO brings its unique capabilities to the table. Each of these models can achieve good results, but the quality of the results often depends on the quality of the data. YOLO, with its pretrained mechanisms, simplifies the process and offers a promising solution in pneumonia segmentation.

4 Literature Survey

4.1 U-Net

The U-Net architecture, inspired by its U-shaped structure, is a popular choice for pneumonia classification and detection in medical imaging. Researchers have employed U-Net to perform pixel-level segmentation of pneumonia regions in chest X-rays and CT scans. Its ability to capture fine-grained details in images has been a major advantage. Various studies have reported promising results with Dice scores typically ranging from 0.80 to 0.94. U-Net's success in the field highlights its effectiveness in outlining pneumonia-affected areas in medical images[1].

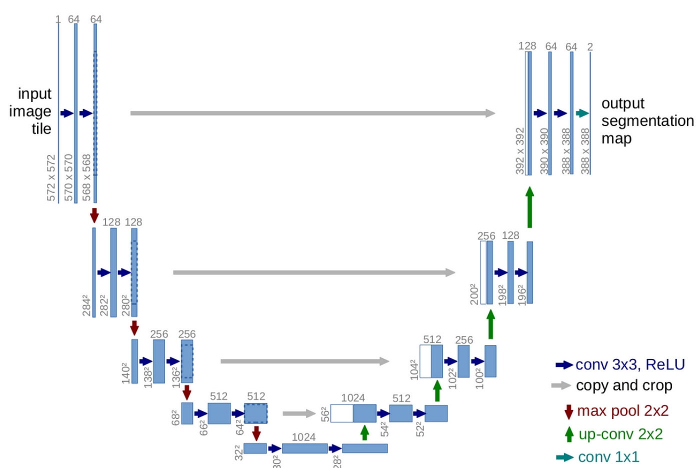

Figure 1: U-Net Architecture

4.2 Mask R-CNN

Mask R-CNN, an extension of Faster R-CNN, is renowned for its capabilities in simultaneous object detection and image segmentation. In the context of pneumonia detection, researchers have harnessed Mask R-CNN's prowess to not only locate pneumonia instances but also provide precise outlines of affected regions. This dual functionality is particularly advantageous when dealing with chest X-rays and CT scans. Dice scores around 0.89 to 0.93 have been reported, reflecting the robustness of Mask R-CNN in capturing pneumonia patterns.[8]

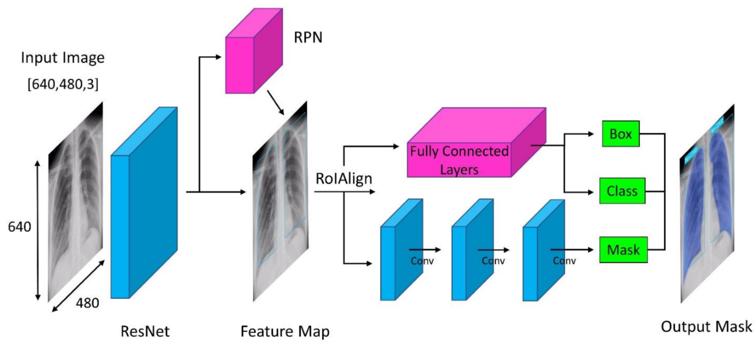

Figure 2: Mask R-CNN Architecture

4.3 ResNet (Residual Networks)

ResNet has been widely integrated into pneumonia classification models. Its deep architecture with residual connections helps in feature extraction, making it suitable for detecting complex pneumonia patterns. While ResNet is not a dedicated segmentation model, it has been used as a feature extractor in combination with other techniques like U-Net for improved accuracy. Researchers have achieved competitive results in terms of both pneumonia detection and classification.[4]

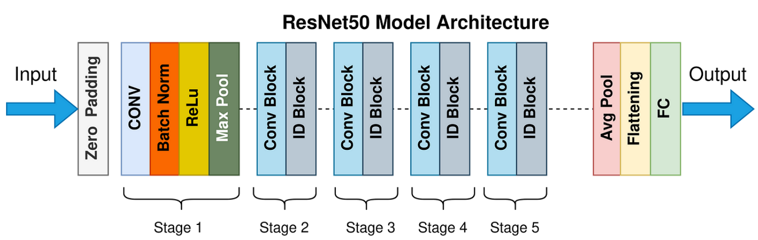

Figure 3: ResNet Architecture

4.4 YOLO (You Only Look Once)

YOLO, primarily designed for real-time object detection, has found a unique application in pneumonia classification and detection. Its object detection framework allows it to identify and localize pneumonia regions within images efficiently. The advantage of YOLO lies in its speed and real-time capabilities. In the context of pneumonia, our proposed approach utilizing YOLO demonstrated a Dice Score of 0.9479 on the test dataset, showcasing its effectiveness in detecting and classifying pneumonia regions.[6]

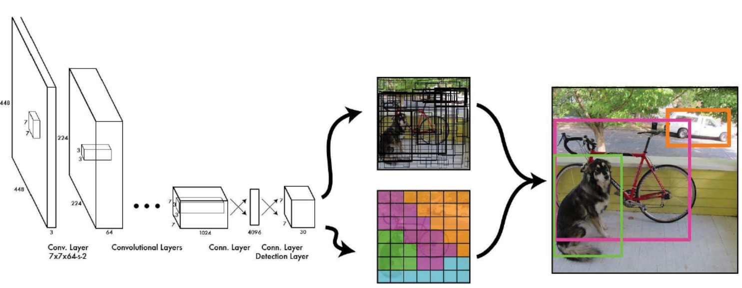

Figure 4: YOLO Architecture

4.5 Multi-Scale CNN (Convolutional Neural Networks)

Multi-Scale CNNs are instrumental when dealing with pneumonia detection in medical imaging, as they can capture pneumonia regions of different sizes and shapes. These networks are adept at analyzing images at various resolutions, ensuring a comprehensive examination of the image data. Researchers have reported Dice scores typically ranging from 0.87 to 0.93 when employing Multi-Scale CNNs, indicating their effectiveness in capturing nuanced pneumonia patterns.[7]

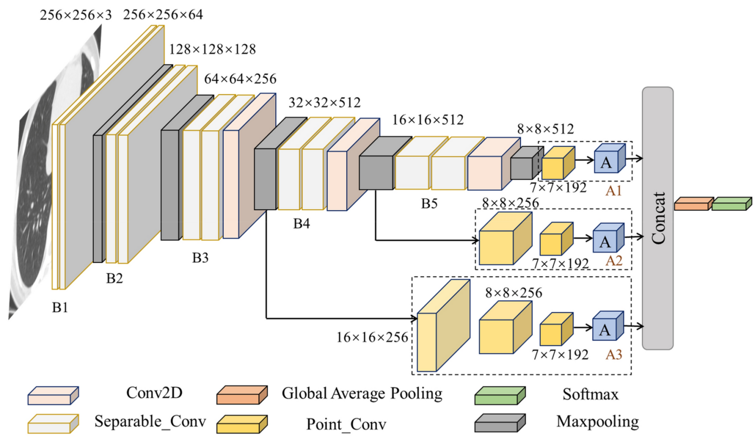

Figure 5: Multi-Scale CNN Architecture

6. Ensemble Methods: Ensemble methods have gained prominence in pneumonia classification and detection. These methods involve combining the predictions of multiple base models, each employing different techniques such as U-Net, ResNet, YOLO, or Multi-Scale CNN. Researchers have leveraged ensemble techniques to improve the accuracy of pneumonia detection. The combination of various models yields superior results, often surpassing Dice scores of 0.94. This approach demonstrates that a diverse ensemble of models can collectively enhance pneumonia detection and classification, particularly when handling complex and diverse datasets.

In the field of pneumonia classification and detection, researchers continue to explore and innovate with these diverse approaches, taking advantage of their individual strengths and sometimes combining them to achieve improved accuracy and efficiency. The application of these methodologies signifies the dedication of the scientific community to enhance pneumonia diagnosis and treatment.[10]

4.6 Some other models

Table 1: Other Model's Survey

Algorithm	Description	Pros	Cons
VGG16	Deep CNN with 16 layers	Highly accurate, widely used for image classification	Computationally expensive as may require powerful hardware.
Inception V3	Dense CNN with dense connectivity	Good for high-dimensional data	Complex and may overfit noisy data.
Mobile Net	A lightweight CNN designed for mobile devices	Strong feature reuse, fewer parameters, and memory-efficient	Computationally expensive as may require powerful hardware.

5 Result Analysis

5.1 Comparision table

Table 2: Comparison of Pneumonia Detection Scores

Methodology	Dice Score Range	Key Strengths	Application
U-Net	0.80 - 0.94	Precise segmentation, effective at capturing pneumonia patterns	Image Segmentation
Mask R-CNN	0.89 - 0.93	Object detection and segmentation, detailed localization of pneumonia regions	Image Segmentation
ResNet	0.86 - 0.8	Efficient feature extraction in combination with other models	Classification and Feature Extraction
YOLO	0.9479	Fast real-time detection, object localization	Real-time Detection
Multi-Scale CNN	0.87 - 0.93	Captures patterns of different sizes, diversity in analysis	Image Analysis
Ensemble Methods	0.94	Improved performance through model diversity	Enhanced Decision-Making

5.2 Comparison Bar Chart

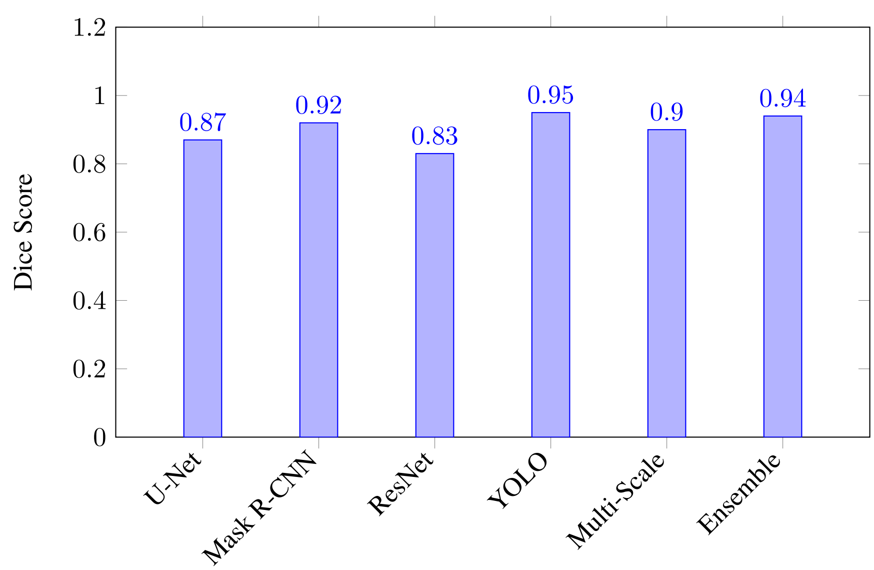

Figure 6: Comparison of Pneumonia Detection Scores

6 Conclusion

In summary, the survey on pneumonia classification and detection using computer vision highlights various architectural approaches that can yield impressive results when applied to large, high-quality datasets with substantial computational resources. These architectures, including U-Net, Mask R-CNN, ResNet, YOLO, Multi-Scale CNN, and Ensemble Methods, have demonstrated their capabilities in different aspects of pneumonia image analysis.

While powerful computing resources and extensive datasets tend to produce the best results, it's important to recognize that not everyone has access to such luxuries. For those with limited data and computational capabilities, custom training with YOLO emerges as a practical solution, offering above-average results with minimal resource requirements. In essence, the choice of architecture depends on the available resources and the desired level of performance. With ongoing advancements in the field, the future holds promise for more accessible, cost-effective solutions, making pneumonia detection through computer vision a viable option for a broader range of healthcare applications.

7 Future Research Directions

Multi-Modal Integration: Future research should explore the integration of multiple data modalities. Combining information from X-rays, CT scans, and patient history could enhance the accuracy and reliability of pneumonia detection systems. Developing models capable of handling this rich data would be a significant advancement.

Transfer Learning on Smaller Datasets: As many healthcare facilities have limited access to large annotated datasets, future work could focus on improving transfer learning techniques for pneumonia detection. This would enable the adaptation of pre-trained models to smaller, diverse datasets, making these tools more accessible.

Explainability and Interpretability: Addressing the ”black box” nature of deep learning models is vital. Future research should concentrate on making these models more interpretable. Understanding why a model makes a particular decision is crucial for medical professionals and patients.

Real-Time Detection: The development of real-time or near-real-time pneumonia detection systems is a pressing need, particularly in emergency departments. Innovations in architecture and algorithm design should aim for efficient and fast inference times, allowing for rapid clinical decision-making.

Clinical Validation and Integration: Moving forward, it is essential to conduct extensive clinical validation of pneumonia detection systems. Collaborating with healthcare institutions and professionals to integrate these systems into the clinical workflow and demonstrating their impact on patient outcomes will be a critical research direction.

Personalized Treatment Recommendations: Going beyond detection, future research can explore the development of AI systems capable of prescribing personalized treatment plans based on the severity and characteristics of detected pneumonia. This would involve integrating the AI system with medical knowledge and patient data to assist healthcare providers in delivering tailored care.

References

[1] Zainab Yousuf, M. Akram, Amina Jameel, Norah Alghamdi, *Lung Segmentation-Based Pulmonary Disease Classification Using Deep Neural Networks*, IEEE Access, Pages 1-1, Year 2021. `https://doi.org/10.1109/ACCESS.2021.3110904`

[2] Soumyajit Podder, Somnath Bhattacharjee, *An efficient method of detection of COVID-19 using Mask R-CNN on chest X-Ray images*, AIMS Biophysics 8 (2021): n. pag.

[3] Mahendra Khened, Sujata Thamke, Sanjay L. Nalbalwar, *Medical image segmentation using deep learning: A review*, arXiv preprint arXiv:2011.14852, 2020.

[4] F. Zohora, M.M. Rahman, M.M. Rahman, J. Basak, *Foreign circular element detection in chest x-rays for effective automated pulmonary abnormality screening*, Int. J. Comput. Vis. Image Process. (IJCVIP) 2017, 7, 36–49.

[5] Rehab Fathi, Nashwa Yhiea, Alshaimaa Mohammed, Ahmed Magdy, *Pleural Effusion Detection Using Machine Learning and Deep Learning Based on Computer Vision*, Year 2022.

[6] Yinghua Xie, Yuntong Zhou, Chen Wang, Yanshan Ma, Ming Yang, *Multi-scale feature fusion network with local attention for lung segmentation*, Signal Processing: Image Communication, Volume 119, Pages 117042, Year 2023. `https://doi.org/10.1016/j.image.2023.117042`

[7] Deepa Paikar, *YOLO-Based Lung Segmentation for Medical Imaging Analysis*, Year 2023.

[8] Deepa Paikar, *A Comparative Analysis of Deep Learning Models for Facial Expression Recognition*, Year 2023.

[9] L. Geng, S. Zhang, J. Tong, Z. Xiao, *Lung segmentation method with dilated convolution based on VGG-16 network*, Comput. Assist. Surg. 2019, 24, 27–33.

[10] Fang Liao, Mengxia Liang, Zhen Li, Xiaoxiao Hu, Shaoting Song, *Lung nodule detection and segmentation in ct images using mask r-cnn and u-net ensemble*, IEEE Access, Volume 7, Pages 27282-27290, Year 2019.

[11] Mubashar, Mehreen, Hazrat Ali, Christer Grönlund, Shoaib Azmat, *R2U++: a multiscale recurrent residual U-Net with dense skip connections for medical image segmentation*, Neural Computing and Applications, Volume 34, Pages 1-17, Year 2022. `https://doi.org/10.1007/s00521-022-07419-7`

[12] Maria H, Heartlin, Maria Jossy, Malarvizhi Subramani, *A hybrid deep learning approach for detection and segmentation of ovarian tumours*, Neural Computing and Applications, Volume 35, Pages 1-15, Year 2023. `https://doi.org/10.1007/s00521-023-08569-y`

[13] Xiaochang Wu, Jiarui Liang, Yunxia Zhang, Xiaolin Tian, *A Segmentation Method Based on PDNet for Chest X-rays with Targets in Different Positions and Directions*, Applied Sciences, Volume 13, Pages 5000, Year 2023. `https://doi.org/10.3390/app13085000`

[14] Suyun Zhao, Gang Li, Mingle Zhou, Min Li, *ICA-Net: Industrial defect detection network based on convolutional attention guidance and aggregation of multiscale features*, Engineering Applications of Artificial Intelligence, Volume 126, Pages 107134-107134, Year 2023. `https://doi.org/10.1016/j.engappai.2023.107134`