narendra

Delhi_Technological_University_Thesis.pdf

Indian Institute of Technology Jodhpur

Document Details

9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

78 Not Cited or Quoted 8%
Matches with neither in-text citation nor quotation marks

6 Missing Quotations 1%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

4% Internet sources
5% Publications
4% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **78** Not Cited or Quoted 8%
  Matches with neither in-text citation nor quotation marks
- **6** Missing Quotations 1%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

4% Internet sources
5% Publications
4% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	**www.mdpi.com**	<1%
2	Publication	**Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Re...**	<1%
3	Publication	**Gaokun Wang, Zhijie Zhang, Wuliang Yin, Guangyu Zhou, Chaofan He, Haoze Che...**	<1%
4	Publication	**Priyanka Anup Ujjainkar, Shital A. Raut. "Enhanced deep learning framework for r...**	<1%
5	Publication	**"Intelligent and Fuzzy Systems", Springer Science and Business Media LLC, 2025**	<1%
6	Internet	**repository.iutoic-dhaka.edu**	<1%
7	Publication	**Lei He, Dasheng Wu, Xinyu Zheng, Fengya Xu, Shangqin Lin, Siyang Wang, Fuchua...**	<1%
8	Internet	**www.algotechniq.com**	<1%
9	Publication	**Rita Aitelhaj, Badr-Eddine Benelmostafa, Hicham Medromi. "APF-YOLOV8: Enhanc...**	<1%
10	Submitted works	**IUBH - Internationale Hochschule Bad Honnef-Bonn on 2026-06-07**	<1%

11 Internet

arxiv.org <1%

12 Internet

cdn.techscience.cn <1%

13 Internet

www.techscience.com <1%

14 Submitted works

Australian National University on 2024-03-26 <1%

15 Internet

www.arabnews.com <1%

16 Submitted works

KICT on 2026-06-07 <1%

17 Submitted works

Universiti Teknologi Malaysia on 2024-03-07 <1%

18 Internet

dspace.iiti.ac.in <1%

19 Internet

www.themoonlight.io <1%

20 Publication

Mohana Karthiga Pasumponthevar, Pandia Rajan Jeyaraj. "Reliability analysis and... <1%

21 Internet

arcenturf.com <1%

22 Internet

ijirt.org <1%

23 Internet

repository.itpln.ac.id <1%

24 Publication

Hamed Amini Amirkolaee, Hossein Arefi. "CNN-based estimation of pre- and post-... <1%

25 Publication

Khang Nguyen, Nhat-Thanh Huynh, Duc-Thanh Le, Dien-Thuc Huynh et al. "A com... <1%

26 Internet

journal.ncit.edu.np <1%

27 Internet

www.unirepository.svkri.uniri.hr <1%

28 Submitted works

Arab Academy for Science, Technology & Maritime Transport CAIRO on 2024-06-27 <1%

29 Submitted works

Asia Pacific University College of Technology and Innovation (UCTI) on 2026-01-02 <1%

30 Publication

B. Saranya Devi, Rimjhim Padam Singh, Angelina George, S. Anirudh, P. Radha Nis... <1%

31 Submitted works

B.V. B College of Engineering and Technology, Hubli on 2025-02-14 <1%

32 Submitted works

Monash University on 2025-10-25 <1%

33 Submitted works

Polytechnic of Namibia on 2023-11-15 <1%

34 Publication

"Advances in Neural Computation, Machine Learning, and Cognitive Research IX"... <1%

35 Submitted works

Anna University on 2025-08-25 <1%

36 Submitted works

Asia Pacific University College of Technology and Innovation (UCTI) on 2023-10-08 <1%

37 Submitted works

Birla Institute of Technology on 2015-02-15 <1%

38 Publication

Chengzhi Deng, Yingbo Wu, Zhaoming Wu, Weiwei Zhou, You Zhang, Xiaowei Sun,... <1%

39	Publication	
Cui-jin Li, Zhong Qu, Shi-yan Wang, Kang-hua Bao, Sheng-ye Wang. "A Method of ...		<1%
40	Submitted works	
Nanyang Technological University on 2026-04-08		<1%
41	Publication	
Paolo Ferro, Harinadh Vemanaboina, Chander Prakash. "Computational Techniqu...		<1%
42	Submitted works	
University of Mauritius on 2026-03-10		<1%
43	Publication	
Zamanidoost, Yadollah. "Early-Stage Lung Cancer Detection Using Deep Learning ...		<1%
44	Internet	
ijsred.com		<1%
45	Internet	
thesai.org		<1%
46	Internet	
www.acadlore.com		<1%
47	Internet	
www.frontiersin.org		<1%
48	Internet	
www.hindawi.com		<1%
49	Internet	
www.ijcs.net		<1%
50	Internet	
www.nature.com		<1%
51	Publication	
"Advanced Intelligent Computing Technology and Applications", Springer Science...		<1%
52	Submitted works	
Monash University on 2020-06-07		<1%

53 Publication

Twahir Kiobya, Junfeng Zhou, Baraka Maiseli, Maqbool Khan. "Attentive context a... <1%

54 Submitted works

Universiti Sains Malaysia on 2024-02-08 <1%

55 Publication

Xueqian Li, Huapan Xiao, Heng Wu, Shaojuan Luo. "Robust steel surface defect de... <1%

56 Publication

Yingqiang Hou, Xindong Zhang. "A lightweight and high-accuracy framework for ... <1%

57 Publication

Junpeng Xu, Lei Shi, Wenqiang Yang, Xiangbo Zhu, Jin Li, Yu Wang. "YOLO-CMDS: ... <1%

58 Submitted works

National College of Ireland on 2022-08-12 <1%

59 Publication

Runwei Ding, Linhui Dai, Guangpeng Li, Hong Liu. "TDD-net: a tiny defect detectio... <1%

60 Submitted works

University of Malaya on 2024-05-28 <1%

61 Publication

Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jac... <1%

Abstract

Printed Circuit Boards (PCBs) are the backbone of modern electronic devices and quality of PCB is paramount for reliability of the system. Traditional inspection techniques for PCB defects by human hand are generally slow, laborious and not very precise in the case of defects which are very small and intricate.This paper provides a framework for automated inspection system for PCBs based on a YOLO (You Only Look Once) architecture which is coupled with a Single Head Self-Attention (SHSA) mechanism to boost the feature representation and inspection ability. The aim is to detect various type of PCB defects using the combination of real-time detection by YOLO and attention based reasoning provided by SHSA to enhance the importance of feature regions and minimize redundancy from backgrounds. Attention mechanism learns the feature space of defect categories by paying special attention to the significant features while disregarding backgrounds in image thereby making the detection of defects, with different sizes, shapes, and intensity variations, robust. The system analyzes PCB image, extracts distinctive features and predicts bounding box for defect regions. This framework is composed of stages such as data preprocessing, feature extraction, multi scale feature fusion, attention enhancement, and detection of defects. These defect classes in PCBs are annotated to train and test the deep learning based object detection system. The experimental validation has been done by reporting various object detection metrics like Precision, Recall, mean Average Precision (mAP) and inference speed. From the empirical results, it is evident that attention mechanisms are beneficial for capturing intricate features and enhancing defect detection accuracy in PCBs. The present work highlights the significance of attention based feature enrichment along with real time object detection technique for industrial automation. The proposed system is extensible and feasible for automating the industrial inspection.

CHAPTER 1

INTRODUCTION

1.1 Overview

Printed Circuit Board (PCB) is a fundamental block of building for our modern electronics. They provide a mechanical support to our components along with a platform inorder to have an electrical connection via the copper tracks present on them. The PCB is a non-conductive substrate which is made up of FR4 fibereglass epoxy laminate. On this non conductive part the copper layers are being etched so that they form the conductive tracks, conductive pads and vias used for the connection of components. This PCB helps us to have an single compact and planar structure used for the connection of components without any need for the hardwire wiring, increasing the reliability and performance of the system. The PCB can be classified into different categories based on the number of conductive layers which they are having along with the mechanical behavior of them. The different kinds of PCB are Single Sided PCB or Single Layered PCB, Double sided PCB or Double layered PCB and Multilayer PCB. Single layered PCB are the PCB where we have the conductive part only on one side of the PCB. On this single side, we would be having our copper traces along with the pads which are being used for the connectivity ofo the compnents. These single sided ones are the best suitable for the low cost and simple circuits which have less power supplies and control boards being present on them. Then comes the Double sided PCB, these are the PCB where we have conductive copper on both sides of the PCB. As there are copper layers on both sides we would be having large copper tracks and pads availability, these copper tracks present on the two different sides are being connected using the plated through holes called as the vias. This double layered PCB would help for the higher density of component mounting with a complex routing, making it suitable for the complex circuits. These are being mostly being used in the consumer electronics and industrial controls because of their ability to handle complex systems compared to the single layered PCB. Then comes the Multilayer PCB, these are the PCB which have more than two copper layers present in them. Here the multiple copper layers are being stacked upon each other and an insulating material is being used between each of these layers providing a insulation between them to avoid any short circuit paths to be formed internally. As there are multiple layers of copper we would have larger number of tracks compared to the previous PCB types, increasing the availability of the copper tracks and vias by a great extend making these PCB highly suitable for the high speed digital system, telecommunication system equipment and advance computing platforms. Here the mulpile layers are being connected using the vias itself, just like the double layered PCB but the only difference is that these are highly complex and need a very high precision while manufacturing. Each type of the PCB has a trade off between the mechanical stability, routing density, speed, complexity and cost. Like in case of the single layered PCB we have high mechanical stability with less routing density making it suitable for least cost and low speed systems. In case of the Double layered PCB we have the good mechanical stability with increase in routing density, which would increasing the cost and speed of the systems for which we are using them. Coming to the multi layered PCB based on the number of the layers we are using , the routing density increasing significantly which increases the speed and cost of the PCB, but also has the issues of congestion and cross talk which should be handled carefully while manufacturing them.

The PCB has conductive paths inorder to have a physical connectivity among all the components which are being mounted on the PCB, so it is important to see that there are no defects present on the PCB as these defects could lead to the malfunctioning of our system causing us a huge loss. Here in the PCB we have multiple surface levels of defects such as the Open circuit, Short circuit, Spur, Spurious copper, Missing hole and Mouse bite. In the Open circuit defect we are having a discontinuity or the breakage on the copper tracks or pads or the vias in our PCB which would lead to a defective conductive path where the conduction doesn't occur and lead to blockage of the current flow. The Open circuit can occur because of no proper completion of etching, mechanical damages which would occurring during the handing of a PCB, because of partial removal of the copper layer during the manufacturing. This causes of an non operational track or part in our circuit lead to no proper availability of the power to all the components which are being connected to that conductive part or section. Then comes Short circuit which is an opposite of the Open circuit, here there is an unwanted path being formed between two copper tracks unintentionally which would lead to overheating, abnormal behaviour along with a latch up condition occurring if we are using a digital circuit and even sometimes cause a permanent damage to our circuit. The Short circuit occurs because of excess copper left during the tracks formation or during the soldering of two closely spaced traced or pads or vias, these can also be caused because of the incomplete etching of the tracks making a bridge between different tracks. Spur is another kind of the surface level defect in the PCB which is being formed because of the unwanted small extension of the copper from the copper tracks or pads. These Spur could lead to formation of the short circuit paths under high voltages or high current situations which is dangerous. These occur because of the over etching during the tracks formation or because of some irregularities occuring in lithography during the manufacturing of PCB or chemical etching. Then we have the Spurious Copper where we have a small part of the unintended track being laid in between the two different tracks in a completely unwanted area and this track is also not being connected to any other track. The would create a parasitic conductive path leading to unwanted linkage of unrelated nets and cause huge leakage current specially in the high impedance analog or the mixed signal circuits. The spurious copper is being formed because of under etching or contamination with a poor mask alignment during the manufacture of the PCB. Then we have the Mouse bite defect which is exact opposite of the spur, where in the spur we have an outward growth of the copper on the tracks which is unwanted in the mousebite we have an inwards reduction in the copper track. These are being caused because of some imperfect depanelization or the routing operations when they are being made to separate boards, these are also being caused during the handling of the PCB when the PCB are being damaged while using them. The mousebite weakens the mechanical integrity of the board which will cause the internal copper layers to be exposed and will even lead to the performance issues, because when we use them in the high voltage or current systems where more head is produced could deplete the copper layer even further and lead to the formation of the open or cause an heat as the copper layer is suddenly being made narrow where there is an mousebite occuring.The next kind of the defect is the defect is the Missing hole where on the connectivity holes where the attachment or the vias are to be formed is partially or fully not present because of the over etching or misalignment or damage caused during the utilization of the PCB. The missing hole would lead to breakage of the vertical connection between the different layers of the PCB especially in the multilayer PCB which is being used for the high speed circuits, causing a great issue as they form open vias, incomplete nets which is not suitable for the systems and lead to the system failure or reduction in the reliability of our PCB.

The PCB forms a crucial component for the connectivity of the electronic circuits in the industries. It becomes important to maintain the reliability, durability and functionality of circuit on the PCB. This can be achieved when we the PCB used is defect free. As, the defects are small and not generally visible to the human eye at an easy way, it is crucial to detect these small defects present on our PCB. To ensure There is a defect free PCB circuit, proposing a deep learning model which would help us in the detection of the PCB surface level defects at an faster and easier rate. This would lead to the utilization of the PCB in a effective way in industry and identify the defects at a faster rate and reduce the number of issues occuring in the PCB at a much faster and easier rate. In order to achieve this I am using Your Only Look Once (YOLO) model which is a deep learning model. The YOLO is being further modified in order to have an much more efficient model for the detection of the PCB defects. The main aim over here is to have the defects being detected at an efficient and effective manner, which is being acieved by our model.

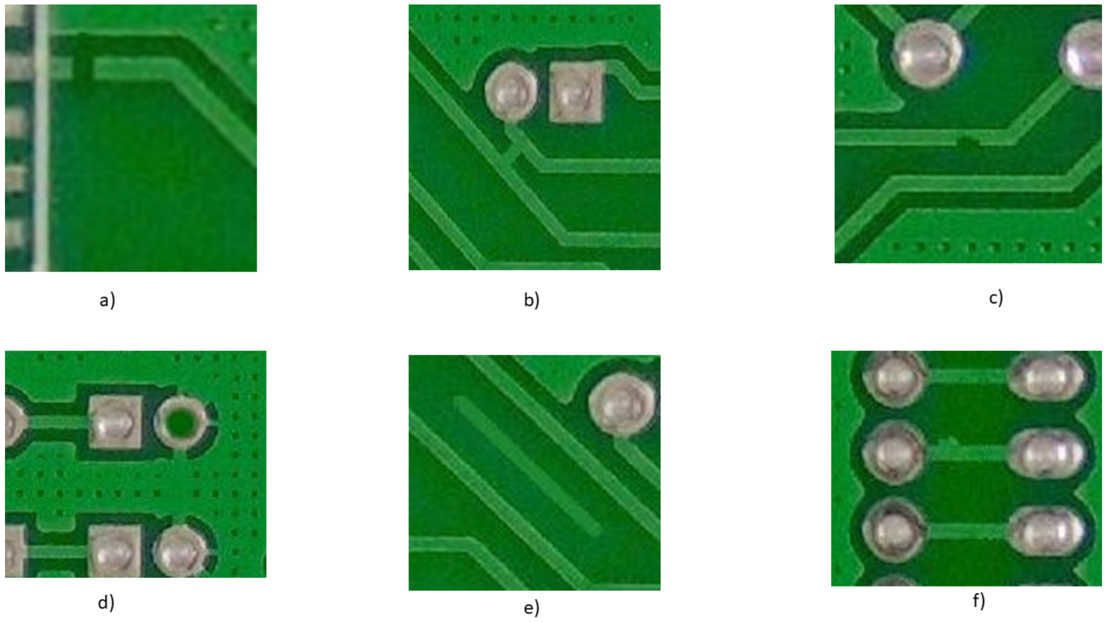

Figure 1.1: Different PCB defects : a) Open circuit b) Short circuit c) Mousebite d) Missing hole e) Spurious copper f) Spur

1.2 Literature Review

The PCB defect detection has become a crucial element for the visual inspection of the PCB for the manufacturing or replacement. This is being done inorder to have an accuracy, consistency along with throughput. To detect these defects some of the authors have developed detection using the RCNN which is a base deep learning model before the usage of the YOLO in order to provide an effective an efficient defect detection model. Faster RCNN is an upgraded version of the RCNN, used by one of the author [1] [2] [3] have introduced a regional proposal networks which is an ened to end trainable two stage detection which has shown a significantly better localization and classification. This has provided a improvement in accuracy and provided a strong baseline. The authors have also implemented a cascade RCNN [4] where they have further refined the two stage detection by introducing a multi stage refinement for the regional defect identification, yielding a high quality of the predictions. As, these CNN models were being used, YOLO has gained the popularity in the industries because of its real time inference with a relatively low computation overhead. There has been multiple versions of the YOLO being developed with time such as the YOLOv5 [5] [6], YOLOv8 [7] [**?**], and YOLOv9 [8] [9] which have been kept on improving the backbone design, its neck architecture and the head for the optimization along with high speed of detection while having the accuracy. These have made YOLO as a good option for the PCB defects inspection where the combination of small scale defects and strict latency requirements is to be maintained.
par In the context of PCBs, many of the authors have adapted the YOLO to handle a

specific challenges for the small scale defects on the PCB. Authors have analyzed a detailed view of how the YOLOv5, YOLOv8 and YOLOv9 would be working giving us an great clarity on the architecture and the improvements which have been occuring with the update of the YOLO. They have clearly demonstrated how the multi resolution has been improved with a lightweight normalization which would help in the PCB defect detection tasks. This has given has a great clarity on how each of these YOLO model is being working helping us have a clear idea on the baseline YOLO models.
par Specialized PCB oriented YOLO models have been made, such at in [?] [10] would integrate the YOLOv8 with the automated PCB defect detection pipelining which would demonstrate us the YOLO based detection is efficient on the multiple defect types while having a good throughput which is suitable for our the production line. In [11] [12] [13], the authors have proposed as YOLO-WWBi which is an optimized YOLO based model for the PCB defect detection in order to enhance the backone, so that there is feature fusion module for better capture of small defects against the complex backgrounds. Here the authors has made a high state of arth model with better mAP for the PCB defect detection. In [12] [2], the authors have made a novel YOLO model called as the YOLO CDMS which is explicitly designed only for the PCB defect detection with a huge emphasis on enhancing the small defect localization with a improved feature extraction and multi scale fusioning. In YOLO DefXpert [13] [14] an other updated YOLO model where the defect detector is enhanced for the better detection of the defects which are being present on the PCB. The authors from [9] [15] [14] have developed TDD Net which is a tiny defect detection network for the PCB where it focus to have a low parameterized design which is suitable for the edge deployment where it clearly shows us the trade off between the model size and defect detection accuracy. The authors from [14] [15] have developed a model known as the DS YOLO which is a small object detection dense model where we would be using an inverted bottle neck with a multi scale fusion network which would share the design principles required for the PCB defect detection. From the [15] [16] where the authors have developed a YOLO model called as the BGF YOLO for the small defect detection in order to have an unmanned aerial vehicle perscpective, through it is for the vehicle it can be seen that it works well for the small object detection so can be even used for the PCB small defect detection if trained properly on the defects. Beyond the pure YOLO models, in the recent times it can be seen that even the CNN transformer models have been combined with the YOLO models and even the attention mechanisms have been used along with the YOLO in order to improve the capturing of better range dependencies and contextual relation between the different parts of the image. The swin transformer by [17] [6] [2] have intoduced a hierarchical transformer design using the shifted windows which would enable the high resolution image understanding without the memory being used much as in case of the full image self attention, this has motivated many of them to work on integrating the transformer based backbones into the YOLO framework. The [18] [19] have even started to use the coordinate attention module into the YOLO as

its backbone component where the information is being sent to channels to guide the feature for selection in lightweight networks, this method has been used in the similar method for the PCB defect inspection to show the relevance of the defect regions while reducing the background noise. The authors from [16] [5] [8] have used the C3TB YOLO where they have integrated a transformer module into the YOLO backbone to improve the feature representation for high resolution remote sensing of the images. In [19] [17] [18] PPLA transformer is being combined along with a YOLO as it is a lightweight transformer which could be used for the linear attention and the pyramidal pooling of the defects which are present. These have clearly shown us a clear trend towards the hybrid CNN transformer based designs where the local feature extraction has been enhanced with global context modeling from the attention mechanisms for the improvement of sensitivity of subtle defects.

par From these it is being clear that YOLO is being actively used in the industry for the small defects detection or the small objects detection and many of them have worked in order to improve the YOLO model into a better and efficient one so that it can be used in a more effective way. It is clear how they have progressively improved the backbone design, neck architecture and the head optimization which would enable the higher detection speed with a god accuracy on the small and densely kept together objects which is similar the the PCB defects. Here I have also developed a new module known as the Shuffle Head Self Attention (SHSA) module. This module is being developed in order to have a lightweight model with self attention present at the neck of the YOLO which would enhance the feature extraction and classification of the defects with a much higher accuracy and precision.

CHAPTER 2

METHODOLOGY

2.1 YOLO Overview

Here we would discuss in detail about the different versions of the YOLO model. YOLO stands for You Only Look Once which is a real time object detection algorithm. Unlike the general detection methods where they perform the localization on a object and would then classify them into the seperate stages, here in YOLO the whole this is being passed through a single forward pass of deep neural networks which would make it significantly faster and a highly effective one for the real time detection and application. The YOLO would generally have three main parts, they are backbone for the feature extraction, neck to have a multi scale level fusion of feature and head for the detection and prediction of bounding boxes. Because of its speed and accuracy the YOLO has emerged as a popular choice for multiple applications like the autonomous driving, for surveillance, industrial inspection and others. Over time the YOLO kept on improving and different versions of the YOLO have been introduced and about them we are going to discuss further in this section.

2.1.1 YOLOv5

YOLOv5 is a fast and efficient real time object detection mechanism which is being developed by ultralytics for the accurate detection and localization of the objects. This is the base for the advances YOLO models. Here in the YOLOv5 from the above Figure we can see that, the left vertical stack is the backbone of the YOLOv5 which extracts features from the input images we would be giving with less resolution and good semantics. Here in the backbone we have Convolution blocks which would do combination of convolution along with batch normalization with the help of activation such as SiLU. Here as we go down the spatial size would keep on shrinking and the number of channels would be growing for deep layers to help them in easier extraction and capturing of the patterns. The C3 modules over here play a key role, these are being derived from the Cross Stage Partial (CSP) network. These blocks would imrpove the feature learning efficiency by dividing the feature maps into two paths, one path would pass through the bottle neck layers while the other path is not being passed through the

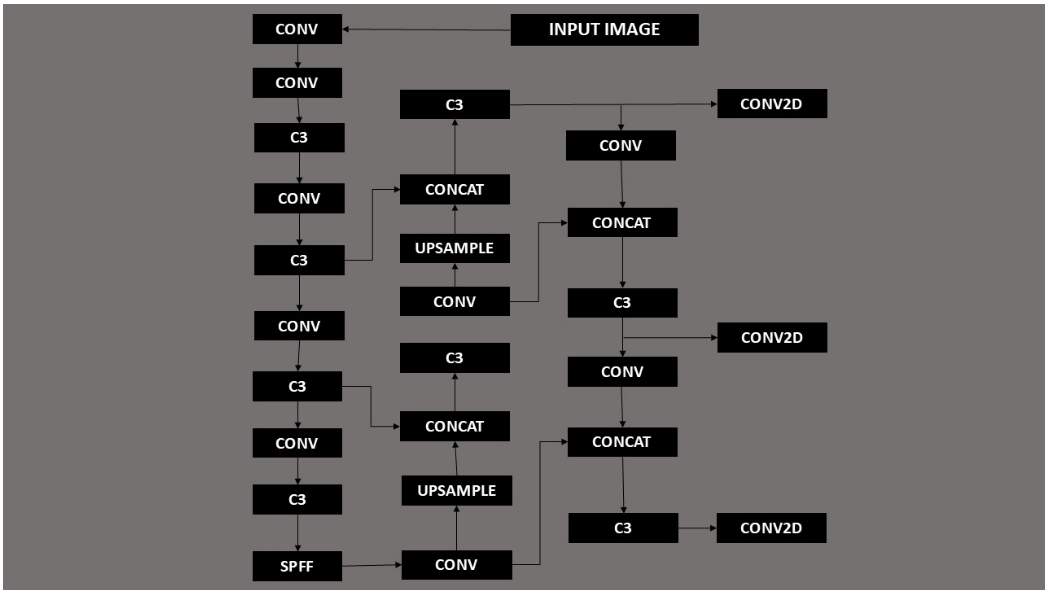

Figure 2.1: YOLOv5 Block diagram

bottle neck and later merges with the other path which would reduce the overall computational redundancy while having a strong gradient flow during the training. This would help in improvement of the stability and detection accuracy without increasing the complexity. At the end of the backbone we use Spatial Pyramid Pooling Fast (SPFF) which is being used for the expansion of receptive field of network, allowing us to capture the information conceptually at different scales. This would help the model to learn both local details as well as the broader details of the images through the two different paths being merged over here, making it learn the different sizes of objects and their values. The middle vertical stack is the neck part of the model, where we do concatenation, convolution and use the C3 modules. Here we combine the different levels of information being obtained from the deeper bottleneck layers which would allow us to have a fine grained spatial information helping us have a better detection of small, medium and large objects present in the image. The Upsampling over here would increase the resolution of deeper feature maps while merging the different feature maps from the backbone. These fused maps from the backbone are being further refined through the C3 modules which would improve the representation quality and reduce noise. Here we have even used PANet based feature aggregation method which would improve the information flow across different values and feature levels which would help in increasing the detection robustness. The detection head is the final stage which is responsible for the generation of the object predictions. Here it would receive the refined features from the multiple scales of neck which is being used for the prediction of bounding box coordinates, confidence score and class probabilities. The multi scale prediction is a crucial feature in this model as it would allow the model to detect objects with different sizes for more effective and high resolution feature maps. The final convolutional layers will convert the information learned from the features into numerical predictions which would correspond to object locations and classification. The features are being further processed using the Non Maximum Suppression (NMS) method to remove the duplicate or low confidence detections, ensuring accuracy.[5]

Thus the YOLOv5 achieves an effective results with a balance between speed accuracy and computational efficiency through its great architecture, making it a widely used solution for the real time object detection tasks.

2.1.2 YOLOv8

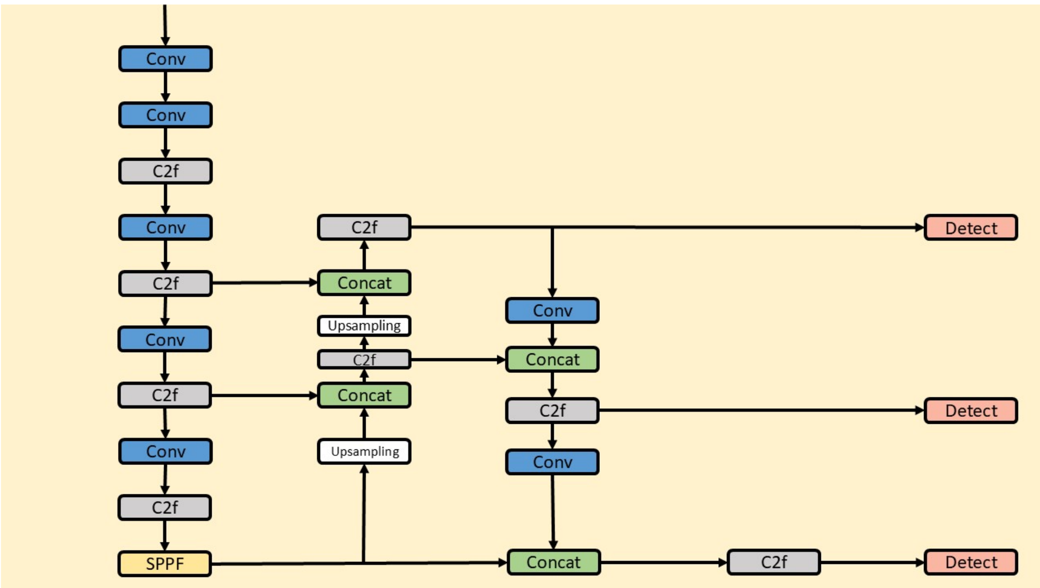

Figure 2.2: YOLOv8 Block diagram

YOLOv8 is a advanced real time object detection model which is being build on the basics of the YOLOv5, designed by Ultralytics. Here in the YOLOv8 we have the same three main parts called as Backbone, Neck and Head for detection. First we would be giving the input to the backbone which is responsible for extracting correct and meaningful information from the input with the help for the Convolution layers, C2f modules athe SPFF. The convolution layers perform feature extraction by applying filters to the input and captuing its edges, textures and shapes, which are used to identify a object. A major upgrade in the YOLOv8 is the usage of Cross Stage Partial Two Convolution Feature Extraction (C2f) which improved the gradient flow with enhancement in feature reuse, reducing the computational overhead compared to previous designs. Unlike YOLOv5 where we use C3 module the C2f allows for a greater information passage by splitting and reusing the features more effectively which would help in improving the learning efficiency and detection accuracy. At the bottom of the backbone we would be using a SPFF just as in the YOLOv5 which is used to increase the receptive aggregation of contextual information at multiple scales, helping us enable the model to better understanding of different sizes without having a increasing in the computational cost. After the backbone we have the Neck part where we would be using Feature Pyramid Network (FPN) along with Path Aggregation Network (PANet) to combine different feature maps, which are being obtained from the different levels of the backbone. Then we would be Upsampling and concatenating them, the upsampling would increase the spatial resolution of deeper features while allowing us to have a fine grained localization details that would have been lost during downsampling. The concatenation would merge high level features from deep layers with a detailed low level features of shallow layers, allowing the model to effectively detect the small and large objects. There are some additional C2f blocks over here which would help in refinement of the fused features, helping in reduction of noise and improvement of quality. This multi scale fusion would significantly improve the object detection performance for small and densely packed objects. Then comes the final part of the model which is the head, used for detection. Here in the YOLOv8 we would be using a achor free detection head which is a major upgrade from the previous YOLOv5 version. The main difference over here is that in the previous version, there was predefined bounding box templates called as anchors used for the prediction while in the YOLOv8 the prediction of the locations id done directly without the usage of the anchors, helping us have a simple training with improved accuracy and efficiency. Here it operates at three different levels for multi scale object detection. Each of the detection layer would predict the bounding box coordinates, confidence scores which indicate the likelihood of object presence and help in class probabilities for the object classification. These raw predictions ar then being sent through the NMS which would remove duplicate overlapping of predictions and retain the most confident detections.[7]

Overall, YOLOv8 is a significant imrpovement for the real time object detection where we have combined different feature extraction with strong gradient propagation along with anchor free detection which enhance the multi scale feature fusion. The optimized architecture gave us an great balance between speed, accuracy and computational efficiency, making it highly effective for the real time applications such as automation, industrial inspection and others.

2.1.3 YOLOv9

YOLOv9 is the next upated version of the YOLOv8. It has introduced several architectural innovation which are being aimed at improving the feature extraction, improving the flow of information and strengthening the multi scale object detection performance, particularly for complex and small scale object detection tasks. Here, similar to the earlier YOLO architectures we have the backbone, neck and head. Here we first use a layer which is used to place hold the tensor operation for managing the input flow without modifying any feature content. This is being followed by a convolution layer which performs the fundamental extraction of feature by applying filters to capture the

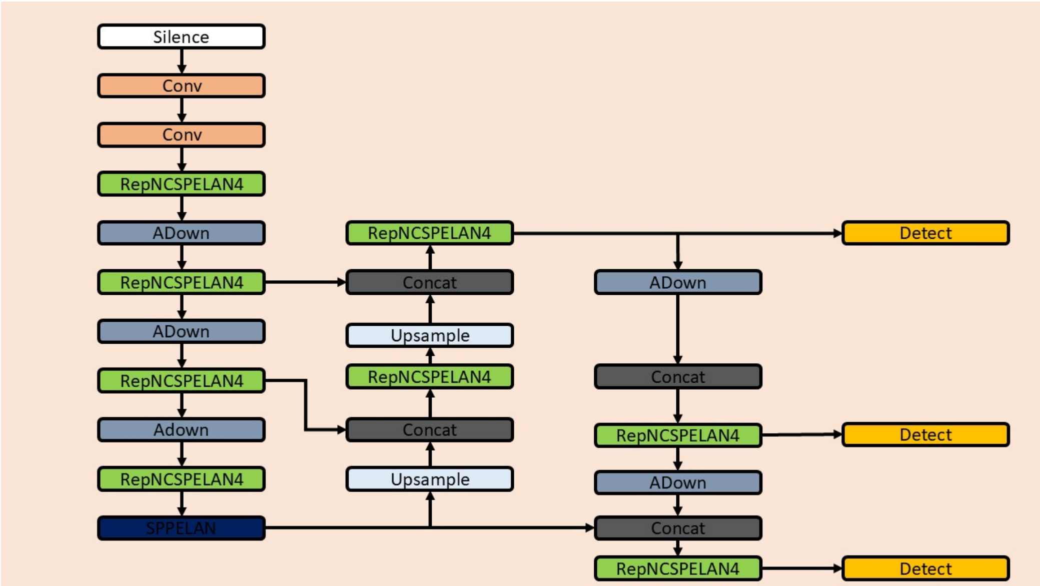

Figure 2.3: YOLOv9 Block diagram

low level visual characteristics like the edges, corners, texture which are used for the pattern or object detection.These convolutional layers also reduce the spatial dimensions while increasing the channel depth, allowing for a deeper semantic representations to be learned. Here in the YOLOv9, usage of the Re parameterized Neural Cross Stage Partial Efficient Layer Aggregation Network (RepNCSPELAN4) is done. This module combines Cross Stage Partial Network (CSPNet) with Efficient Layer Aggregation Network (ELAN) inorder to improve the feature for learning while maintaining the computational efficiency. Here RepNCSPELAN4 enhances gradient propagation, helping in promoting feature reuse and strengthening of representation learning. This helps in getting richer information flow across multiple branches compared to the traditional convolution blocks and this would help in learning depth parameters without causing any instability. Here even we use Average Downsampling (ADown) which performs downsampling while preserving all the important features in an effective way than the conventional convolutional downsampling. The backbone at the end uses the Spatial Pyramid Pooling Efficient Layer Aggregation Network (SPPELAN) which combines multi scale contextual feature extraction with an efficient feature aggregation. This would expand the receptive field of the network which would allwo the neurons to capture both the local and global contextual information and improve the detection of objects while having a varying scales and appearances. Then comes the neck section where we do the multi scale integration of the features for the fusion and refinement. Here we would be using the upsampling, concatenation, ADown and RepNCSPELAN4 modules. the upsamplin would increase the spatial resolution for deeper feature maps and improve the localization of the details without a loss in the information. The concatenation would merge high level information with the low level information for having a rich representation. This fusion would allow the model to detect all kinds of the objects. The usage of the RepNCSPELAN4 in the neck further increases the information by having a refinement and increase in robustness with reduction in redundant information. Then in the final stage which is the head, Programmable Gradient Information (PGI) is being used, this introduces a training oriented strategy to reduce the information bottlenecks and help in improvising optimization. The raw predictions over here are being sent through the NMS in order to eliminate the duplicate overlapping of the detections occuring and retain only the most confident object predictions.[8]

Overall, in YOLOv9 a significant advancement has been made by integrating efficient feature aggregation with improved gradient flow and optimized downsampling, giving us an enhanced multi scale contextual learning. These make it highly effective for real object detection tasks.

2.2 Proposed Methodology

After analyzing the YOLO models, developed a novel called as the YOLO SHSA where SHSA stands for Shuffle Head Self Attention. This model is being being built on the YOLOv8 as base. This is because of the robust model of YOLOv8 which makes it one of the popular choice for all the defect detection in the PCB defects. YOLOv8 can be divided into three main parts, backbone, neck and head part and the details regarding how the YOLOv8 works and functions and works is being explained in the above. SHSA stands for Shuffle Head Self Attention, it is an upgraded form of standard self attention mechanism. It is being designed to enhance the power of representation which is being achieved by using the convolutional networks such as YOLOv8. In a general detection module, the backbone features are being generated by staking of the convolutional blocks that will primarily capture local patterns using their receptive fields. But in Self Attention we will allow each spatial position in the feature map to have a direct interaction with all the other positions, helping us have an long range dependencies and global context. In case of the general convolutional detection module the head structure will split its channel dimensions into multiple subspace inorder to have different subsets of channels, this will help them learn and focus on different types of patterns and the relationship present in between them. The Shuffle Head comes into picture over here, it mixes outputs of these heads on the channel dimensions so that they capture the information of distinct heads and redistribute those back so that they can be jointly exploited by all the subsequent layers.

To clearly understand how the Shuffle Head Self Attention would work, let us take a input Feature map X of size NxC, where N indicates the number of the spatial locations and C is the number of the channels present. The Shuffle Head Self Attention would first dive the C channels into H heads of equal size, such that C=Hxd, where d is the channel dimension for each head. For a head h, the corresponding slice of the input is

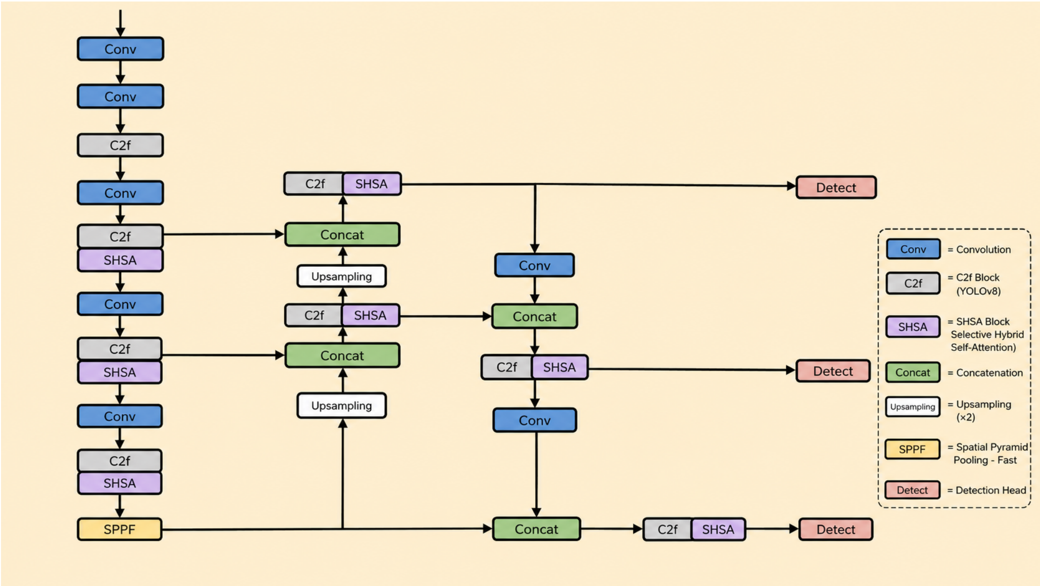

Figure 2.4: YOLO-SHSA Block diagram

denoted by formula. From this we would we obtaining the Query(Q_h), key (K_h) and value(V_h) for each of the h. These query, key and value are being calculated using the below equations 1,2 and 3

$$Q_h = X_h W_Q \tag{2.1}$$

$$K_h = X_h W_K \tag{2.2}$$

$$V_h = X_h W_V \tag{2.3}$$

In the above equations, W_Q, W_K and W_V represent the weights which are being associated with the matrices. The Self Attention response of head his is then being computed using the scaled dot product formulation as shown in equation 4

$$A_h = softmax\left(\frac{Q_h K_h^\top}{\sqrt{d}}\right) \tag{2.4}$$

From the above we have calculated A_h,the attention matrix for that head whose entries represent how strongly each position would attend to the other positions available. The division by root d stabilizes the magnitude of the dot products and softmax function is being applied row wise so that A_h forms a probability of distribution. The output O_h is an weighted combination of value vectors where the learned attention scores would be giving us the weights. Once the O_h is being calculated for all the heads they are all being concatenated to reconstruct a full feature map as shown in equation 5

$$O_h = A_h V_h \tag{2.5}$$

$$O = [O_1, O_2, \dots, O_H] \in R^{N \times C} \tag{2.6}$$

From Equation 6 above it is clear that O is a group of heads which is combination of all the O_h, meaning that the channels occupying contiguous ranges would correspond to the same attention subspace. The Shuffle operation is being implemented over here in order to break the grouping and see that there is good interaction among the different heads so that the information learned by the different heads is being shared with each other and the shuffling can be represented in equqtion 7.

$$\tilde{O} = Shuffle(O) \tag{2.7}$$

Because of the shuffling the subsequent convolutional or feed forward layers would receive feature vectors in which each local neighborhood of channels would combine the information from different attention heads rather than focusing only on a single head features. Finally giving rise to final output of Shuffle Head Self Attention with W_O being the learned weight matrix and Y being final output as shown below in equation 8

$$Y = \tilde{O} W_O \tag{2.8}$$

So, the above is being integrated into the YOLOv8 as this will serve as a great complementary part to our convolutional backbone and decoupled detection head present in YOLOv8. The YOLOv8 which has convolution and C2f blocks present in its backbone is an effective one for extracting the local and hierarchical features but is limited in capturing the long-range dependencies which the object appearance and context are far apart spatially. By the usage of the Shuffle Head Self Attention in the deeper stages of the backbone or neck where we will be having a lower spatial resolution with high features being available, we will be able to have a good information accumulation from all the different positions of the feature map. The multi head design will make different heads to focus on different features across the data and during the shuffling step, we will be exchanging the information from all of the heads and get higher feature values. This ensures that the contextual information modeled by different heads are being available to all of the convolutional heads which will help in the detection task where we have subtle defects which must be recognized and help in easy classification with localization performance. As the attention is being computed on an reduced per head dimension and we are applying the shuffling which is a simple permutation, the addition of Shuffle Head Self Attention will not add a lot of overhead computation for our YOLOv8 making it a great choice for the improvement of the detection of classification in YOLOv8.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Training and Experimental Setup

The experiment is conducted using PCB defect dataset which is available to us in a open manner[9]. Here the different PCB defects have been given to us, which are Open circuit, Missing hole, short circuit, Spurious copper, Mousebite and spur. Each defect type is a manufacturing issue that will cause significant effect on the functionality and reliability of the PCB. Here in the dataset we have 10,668 annotated images where we have the annotation present in the yaml file with respect to each of the image highlighting to us where the different defects are being present, which would be used for the training of the model. To have an unbiased evaluation, the dataset is being divided into train, test and validation in the ratio of 80:10:10. So, there are total of 8,534 images in the train, 1,066 images for validation and 10,068 images for test. The training set is being used to optimize the model, validation is used to monitor how the model is being performing after learning the parameters and test is used to finally evaluate a unbiased assessment of trained model. This whole experimental setup is being implemented on the Google colab which provides a access to cloud based GPU, suitable for intense and deep learning experiments. The model training is being performed on the T4 GPU which is a NVIDIA tuning architecture and provides us with an 16GB of GPU memory. The Tesla T4 is widely being used to adopt the macchine learning due to its balance of computational efficiency with inference capability. Here we used the Automatic Mixed Precision (AMP) training to improve the computational performance and reduce memory consumption. The AMP combine half precision and single precision floating point operation, accelerating matrix computations while having the numerical stability. This is beneficial for object detection models involving large convolutional architectures as it would allow faster training without substantial degradation in model accuracy. The AdamW optimizer is used to prevent overfitting while having a stable parameter updates during optimization. The learning rate over here is being 0.01 which provides us with a balance between convergence speed and training stability. With all these we have trained the YOLO SHSA for 50 epochs allowing for a sufficient information for learning meaningful defect representation from the training data.

The total training for YOLO SHSA has taken 4 hours of times which demonstrates the computational efficiency of proposed approach considering the dataset size and model complexity. This training duration was influenced by factors such as dataset size, GPU performance, image resolution, optimization strategy and mixed precision acceleration. For the training I have used the COCO model with the pretrained weights. Here to improve the robustness of defect detection the training is using augmentation strategies integrated with Ultralytics for the training framework. These would assist in improving the generalization by giving the model to have varying spatial transformations and image distributions for training. For the Performance evaluation, used Precision, Recall and Mean Average Precision(mAP). Precision is calculated using confusion matrix of the given model as shown the equation 9, it is to measure the accuracy with which objects are being detected by our model.

13

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{3.1}$$

17

Recall is another standard measurement used to give us information about the ground truth of images, which indicates how many number of actual images are being detected accurately, it is also being calculated using the confusion matrix using the equation 10

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{3.2}$$

33

Mean Average Precision (mAP) is used to let us know the trade off between the precision and recall values. We have even calculated mAP50 which is used to report average precision of detection, here it shows us how many detections boxes are present with atleast 50 percent of overlap with the actual defects. These are being represented in equation 11 and equation 12.

$$\text{AP} = \int_0^1 P_i(R)\, dR \tag{3.3}$$

$$\text{mAP@50} = \frac{1}{N} \sum_{i=1}^{N} \int_0^1 P_i(R)\, dR \tag{3.4}$$

The other metric used over here is the mAP50:95 which is being shown in equation 13, it is being used to tell us how the detection would be working under the different levels of strictness.

$$\text{mAP}_{50\text{–}95} = \frac{1}{10N_c} \sum_{t \in \{0.50, 0.55, \ldots, 0.95\}} \sum_{c=1}^{N_c} \text{AP}_c(t) \tag{3.5}$$

3.2 Result

The obtained confusion matrices of the baseline YOLO and proposed YOLO-SHSA model after training are illustrated in Figure 3.2. The confusion matrices provide a class-wise visualization of the accuracy and misclassification behaviors of prediction results of all PCB defect classes. The Precision-Recall curves of all the evaluated models are illustrated in Figure 3.3. This chart is used to determine the performance of mAP@50. It plots the values for Precision vs. Recall for every value of the confidence threshold from 0.0 to 1.0 and shows the variance of the F1-score to the confidence threshold is represented in Figure 3.4. This analysis is very helpful to understand the consistency of all models under varying required certainty of detection, and in order to determine the required certainty value of prediction for a certain task. The measured performance of precision, recall, mAP@50 and mAP@50:95 of all the compared models is tabulated in Table 3.1. From Table 3.1, it is clearly observed that all the models performed well for PCB defect detection dataset, which indicates high detection precision and recall for all models. The YOLOv5s yields a precision of 0.9433 and recall of 0.9898, which indicates that defects are correctly identified with a small amount of missed detections but the precision is low when the number of false positive is high. The YOLOv8s gives a precision of 0.9563 and the highest recall among the baselines (at 0.9918) for defect detection which shows a high detection sensitivity. The YOLOv9s also yields a high precision (at 0.9533) and the highest recall (at 0.995) which implies that defects are almost completely identified by the model. Comparing with the baseline models, the proposed YOLO-SHSA yields the highest precision (at 0.9686) which is more accurate in terms of prediction reliability compared to other models and also gives an acceptable recall value (at 0.9881). This means that, though the missing detection rate is slightly higher than that of the YOLOv8s and YOLOv9s, the detection accuracy for those detected bounding boxes is higher. In real industrial detection process, missing false positive result is quite important for the sake of reducing the rejection rate. Moreover, the proposed model acquires the highest mAP@50:95 score (at 0.634) which means that the localization accuracy of the predicted boxes under lower IoU thresholds are more superior. It can be inferred that the SHSA enhanced model is better in detecting defects and bounding box prediction.

Table 3.1: Performance comparison of different YOLO models on PCB defects

Model	**Precision**	**Recall**	**mAP@50**	**mAP@50:95**
YOLOv5s	0.9433	0.9898	0.989	0.5771
YOLOv8s	0.9563	0.9918	0.988	0.613
YOLOv9s	0.9533	0.995	0.982	0.630
YOLO-SHSA	0.9686	0.9881	0.988	0.634

As we observe in Figure 3.2, YOLOv5s and YOLOv8s demonstrate acceptable classification accuracy, but high confusion happens when detecting similar defects (e.g.

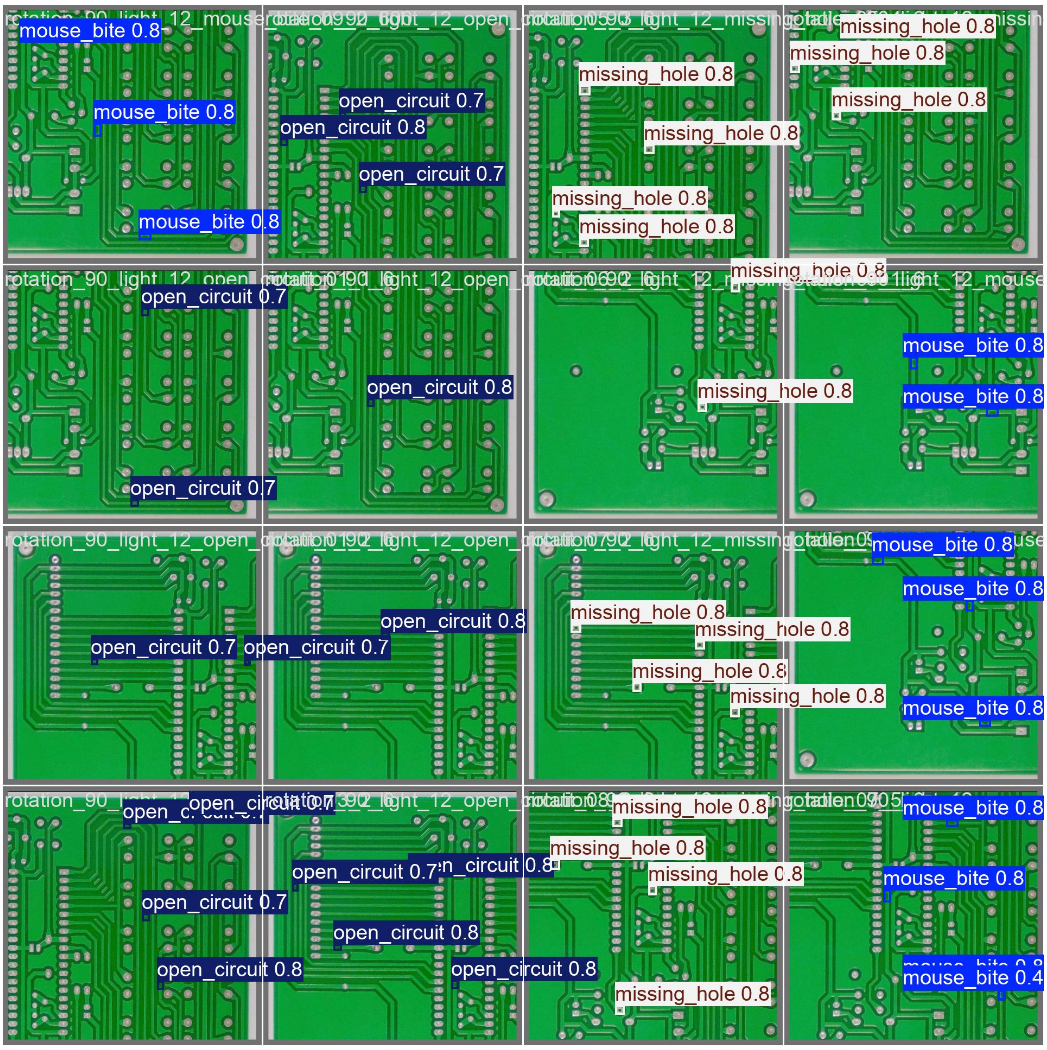

Figure 3.1: YOLO-SHSA predictions

Spur and Spurious Copper). Those defects differ subtly, so they become hard to distinguish. YOLOv9s improves on classification results by reducing some confusion cases between Spur and Spurious Copper, yet there is still some miss classifications of the spurious copper part. YOLO-SHSA, on the other hand, results in less incorrect classifications between different defect types. These improvements of feature discrimination capability with the integration of SHSA could be attributed to its performance on subtle distinction of the defect types with small and different structures. As we see in Figure 3.3, all the models perform well on the task since their PR curves are at the top left. However, the PR curve of YOLO-SHSA for most defects demonstrates more robustness by keeping the precision constant while increasing the recall rate. The constant PR curves are very useful when selecting an operating point. By looking at Figure 3.4, we observe that YOLOv5s and YOLOv8s obtain a high peak of the F1 score but drop fast when confidence threshold increases, which is a sign that they are not robust enough when stricter filtering condition applies. YOLOv9s show a comparatively smoothly

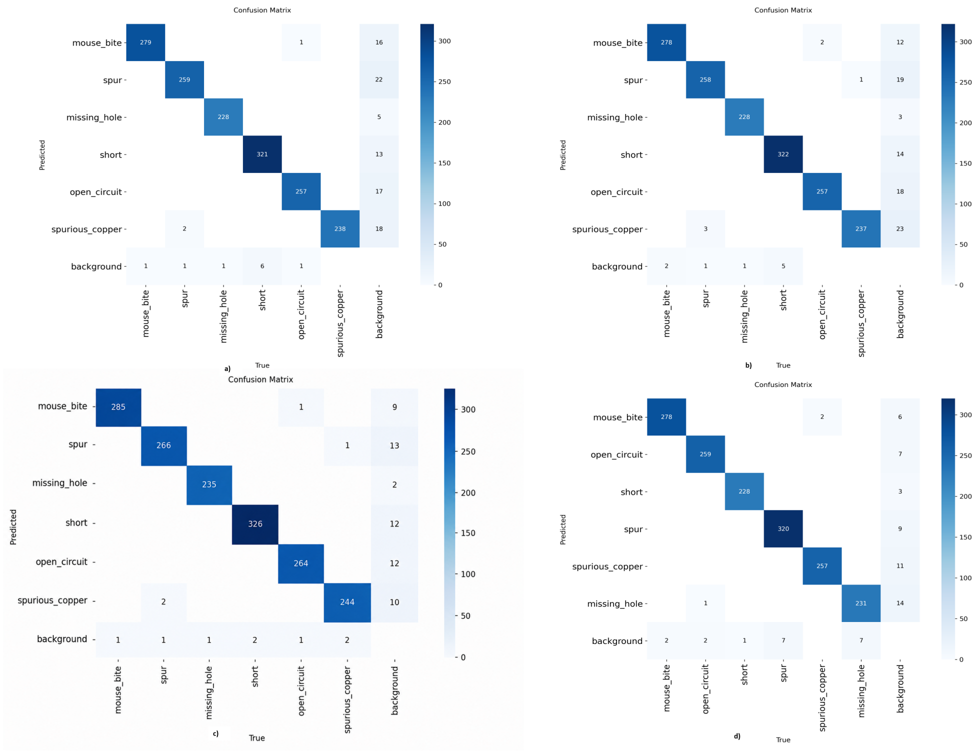

Figure 3.2: Confusion matrices of YOLO models: (a) YOLOv5, (b) YOLOv8, (c) YOLOv9, (d) YOLO-SHSA.

changing F1 curve than the former two, while the proposed YOLO-SHSA has the most robust F1 curve for most cases. The smooth F1 curve illustrates that the proposed model could be more robust under varied confidence threshold settings, indicating that with SHSA added, feature selection is improved in terms of its strength. The detection capability of the proposed model can be seen in Figure 3.1. We can find that almost all the defects on the PCB surface are correctly identified, with clear detection bounding boxes and high detection confidence values (most of them are larger than 0.70). The reasons for improved detection ability are that SHSA can focus more on discriminative feature regions and it is able to extract the finer structural features of defects that are hard to extract with traditional YOLO models, as defects are mostly small and subtle features.

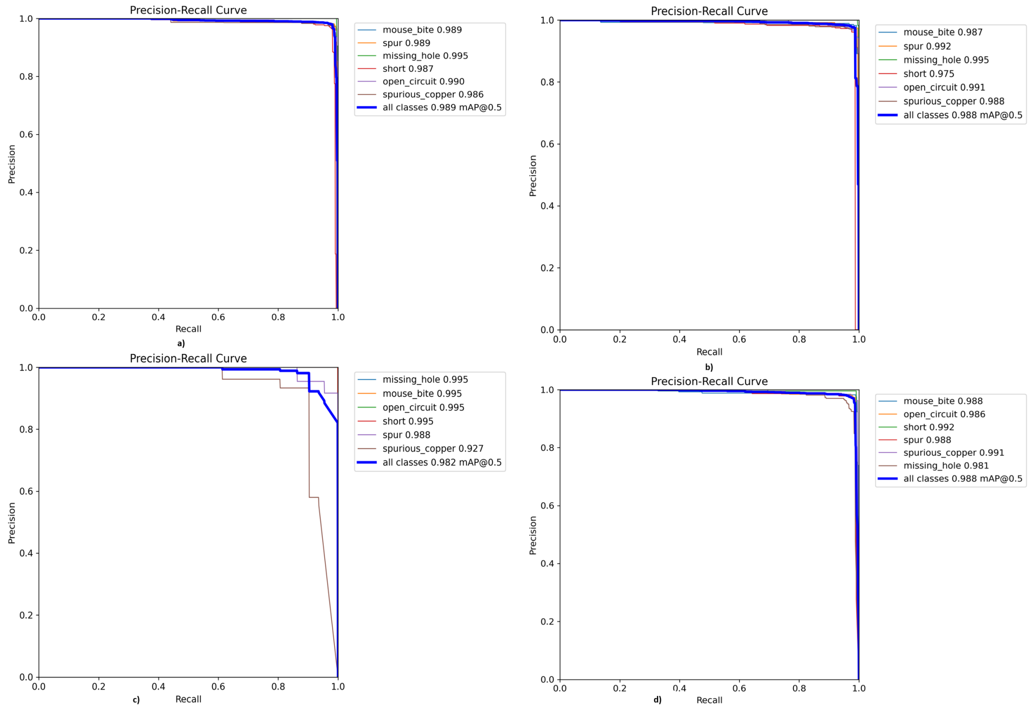

Figure 3.3: Precision vs Recall graph of YOLO models: (a) YOLOv5, (b) YOLOv8, (c) YOLOv9, (d) YOLO-SHSA.

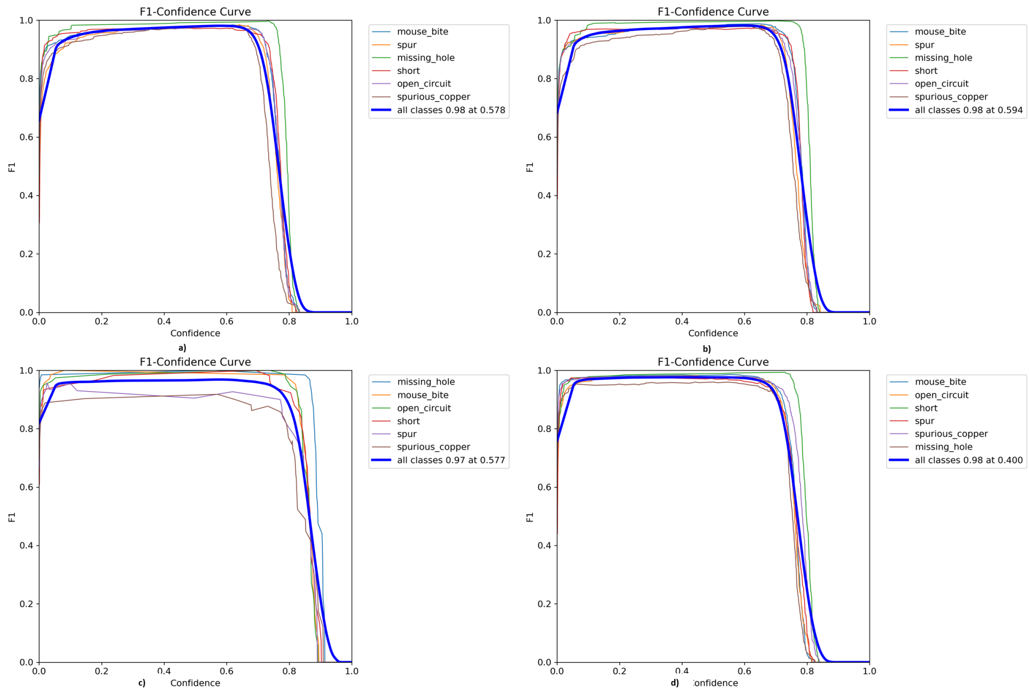

Figure 3.4: F1-curve of YOLO models: (a) YOLOv5, (b) YOLOv8, (c) YOLOv9, (d) YOLO-SHSA.

CHAPTER 4

CONCLUSION AND FUTURE SCOPE

4.1 Conclusion

In this work, a deep learning method for automatic PCB surface defect detection with enhanced YOLO network with Single Head Self Attention (SHSA) is presented. Inspection of printed circuit boards is very important in the modern electronic production where small surface defects can have severe influence on the product reliability, working performance and safety. Manual inspections are slow, inconsistent, error-prone and lack the precision required for identification of fine details of smaller defects. To address these issues this paper aims at developing a rapid and accurate automated defect detection system for identification of various surface defects in printed circuit boards with higher reliability. The system has employed a public PCB defect dataset which contains six main types of defect which are 'Open Circuit', 'Short Circuit', 'Spur', 'Spurious Copper', 'Missing Hole' and 'Mouse Bite'. To evaluate the system performance, this dataset is segregated into training, validation and test subsets. YOLO based object detection architectures i.e. YOLOv5s, YOLOv8s, YOLOv9s and the proposed YOLO-SHSA architecture were implemented and tested on same experimental conditions. The detection performance is analyzed on the basis of standard object detection evaluation parameters i.e. Precision, recall, mAP@50 and mAP@50:95. From the experimental results it is clear that the base YOLO models also exhibit quite a high accuracy in defect detection which confirms the effectiveness of modern object detection models for industrial inspection tasks. However, the YOLO-SHSA model exhibits highest general consistency across all detection metrics. The addition of SHSA attention module in YOLO helps to focus on relevant defect regions effectively, enriching the feature representation of small and complicated defects. In this regard, the proposed model shows the highest precision as compared to all the base YOLO models that can infer that there are fewer number of wrong predictions from the model. Though few base models obtain higher values for recall, the YOLO-SHSA model has a more practical trade-off. In a typical quality control inspection environment, false alarm rate is as important as recall since high rate of false alarms can lead to unnecessary discarding of acceptable components and increased production cost. Improved mAP@50:95 indicates better localization accuracy on tougher threshold. The integrated attention mechanism in feature extraction helps in enhancing this feature and that is why the proposed YOLO-SHSA performs better in case of mAP@50:95. Through the use of confusion matrix, precision-recall curves and F1-confidence curves it is confirmed that the proposed method is efficient enough. It can be seen from both, the confusion matrix of the proposed model and its neighbors, that the class confusion is significantly reduced in the proposed method compared to base models, similarly both precision-recall curves and F1-confidence curves show that the proposed YOLO-SHSA performs significantly better for the detection of PCB defects with lower threshold. Moreover the qualitative results of the detection show accurate localization of many defects on a printed circuit board. In this work, it is demonstrated that integrating lightweight attention modules like SHSA into state-of-the-art real-time object detection architectures enhances their performance for PCB defect detection with a significant reduction in computational complexity. The proposed system provides a practical and effective solution for automated quality inspection in manufacturing by achieving higher detection accuracy, precise localization, and robustness. The results of this work clearly show the potential of attention based deep learning architectures for improved quality control systems in manufacturing industry and for reduction in reliance on manual inspection techniques.

4.2 Future scope

While the proposed YOLO-SHSA model shows promising results in detecting surface defects of PCB boards, there are still many rooms for improvement and extension for this work: improving detection accuracy, computational efficiency, generalization ability, and industrial application, etc. One direction is to study on larger and more diverse PCB defects datasets. The current work only considered one public dataset which contains six types of defects under controllable lighting. But real industrial environments are far more complex, with lighting, noise, variety of design structure, fabrication variations and defects appearance difference, so it should be evaluated on larger industrial-scale datasets that have more diverse defect categories to increase the generalization ability. Another direction to study is exploring different advanced attention mechanisms and feature enhancement methods, though SHSA has obtained improvements of defect representation and localization performance. Some possible mechanisms including Triplet Attention, CBAM, Coordinate Attention, Transformer based self-attention, hybrid spatial-channel attention and so on, should be discussed or a comparison experiment on several lightweight attention modules should be implemented to find a better one for the PCB inspection application. For the PCB inspection task, it often involves detecting very tiny structural abnormality which occupied a small image region, so optimizing small defect detection may be another worthy research direction. Some approaches such as super-resolution preprocess, multi-scale feature fusion, adaptive anchor optimization and small object detection network could be used to improve the performance for small defect inspection. Computation efficiency can be improved as well to achieve practical industrial application. The proposed work had achieved a practical performance on GPU which is suitable for cloud environment. In real manufacturing environments, some resource-limited edge devices such as embedded hardware may be required for real-time inspection. The model can be further optimized through methods like model pruning, quantization and knowledge distillation to reduce memory requirements and inference latency while maintaining high accuracy. Combining the proposed defect detection framework with real-time industrial inspection systems is another significant aspect to be studied in the future. This would allow continuous monitoring of PCBs using live camera feeds, automated conveyor inspection machines and manufacturing quality control pipeline. The combined system can be used for real-time PCB inspection, instantaneous identification of defects and automatic rejection of defective components which could improve production efficiency. Additionally, multi-modal inspection that combines RGB image-based defect detection with thermal imaging, X-ray imaging or infrared detection methods can be utilized to enhance detection performance for invisible structural defects. Furthermore, investigating beyond detection toward defect severity assessment and prioritizing defect detection can extend the application scope of the proposed framework and provide more useful information for decision making in industry. In conclusion, YOLO-SHSA framework showed great promise on PCB surface defects detection; however, more efforts are needed for better generalization, efficiency, scalability, industrial applicability and expandability.

REFERENCES

[1] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[2] Y. Cheng, Y. Cao, H. Yao, W. Luo, C. Jiang, H. Zhang, and W. Shen, "A comprehensive survey for real-world industrial surface defect detection: Challenges, approaches, and prospects," *Journal of Manufacturing Systems*, vol. 84, pp. 152–172, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612525002845

[3] Y. Leng and J. Liu, "Improved faster r-cnn for steel surface defect detection in industrial quality control," *Scientific Reports*, vol. 15, no. 1, p. 30093, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-12740-x

[4] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," 2017. [Online]. Available: https://arxiv.org/abs/1712.00726

[5] R. Khanam and M. Hussain, "What is yolov5: A deep look into the internal features of the popular object detector," 2024. [Online]. Available: https://arxiv.org/abs/2407.20892

[6] J. Doherty, B. Gardiner, E. Kerr, and N. Siddique, "Bifpn-yolo: One-stage object detection integrating bi-directional feature pyramid networks," *Pattern Recognition*, vol. 160, p. 111209, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320324009609

[7] M. Yaseen, "What is yolov8: An in-depth exploration of the internal features of the next-generation object detector," 2024. [Online]. Available: https://arxiv.org/abs/2408.15857

[8] ——, "What is yolov9: An in-depth exploration of the internal features of the next-generation object detector," 2024. [Online]. Available: https://arxiv.org/abs/2409.07813

[9] R. Ding, L. Dai, G. Li, and H. Liu, "Tdd-net: A tiny defect detection network for printed circuit boards," *CAAI Transactions on Intelligence Technology*, vol. 4, 05 2019.

[10] A. Rahimi, M. Anvaripour, and K. Hayat, "Object detection using deep learning in a manufacturing plant to improve manual inspection," in *2021 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 2021, pp. 1–7.

[11] Y. Zhao and Z. Jiang, "Yolo-wwbi: An optimized yolo11 algorithm for pcb defect detection," *IEEE Access*, vol. 13, pp. 74 288–74 297, 2025.

[12] J. Xu, L. Shi, W. Yang, X. Zhu, J. Li, and Y. Wang, "Yolo-cmds: An enhanced method for pcb defect detection," *Integration*, vol. 109, p. 102732, 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167926026000878

[13] P. Selvam, R. Rajasekar, C. Gunasundari, S. Janu Priya, M. Murugappan, and M. E. H. Chowdhury, "Yolo-defxpert: An advanced defect detection on pcb surfaces using improved yolov11 algorithm," *IEEE Access*, vol. 13, pp. 143 085–143 101, 2025.

[14] H. Zhang, G. Li, D. Wan, Z. Wang, J. Dong, S. Lin, L. Deng, and H. Liu, "Ds-yolo: A dense small object detection algorithm based on inverted bottleneck and multi-scale fusion network," *Biomimetic Intelligence and Robotics*, vol. 4, no. 4, p. 100190, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667379724000482

[15] J. Mei and W. Zhu, "Bgf-yolov10: Small object detection algorithm from unmanned aerial vehicle perspective based on improved yolov10," *Sensors*, vol. 24, no. 21, p. 6911, 2024. [Online]. Available: https://doi.org/10.3390/s24216911

[16] Q. Wu, Y. Li, W. Huang, Q. Chen, and Y. Wu, "C3tb-yolov5: integrated yolov5 with transformer for object detection in high-resolution remote sensing images," *International Journal of Remote Sensing*, vol. 45, pp. 2622–2650, 04 2024.

[17] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," 10 2021, pp. 9992–10 002.

[18] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," 2021. [Online]. Available: https://arxiv.org/abs/2103.02907

[19] X. Song, Y. Tian, H. Liu, L. Wang, and J. Niu, "Ppla-transformer: An efficient transformer for defect detection with linear attention based on pyramid pooling," *Sensors*, vol. 25, p. 828, 01 2025.