NetInspect: Exploring CNN based Deep Learning models for Defect Detection from images of PVC specimens and Heat Sinks.

REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF

BACHELOR OF TECHNOLOGY
IN
INFORMATION TECHNOLOGY/ COMPUTER SCIENCE & ENGINEERING

By

Sanjana Das
Roll Number: 200102016
&
Avimanyu Dutta
Roll Number: 200102019

UNDER THE GUIDANCE

OF

Dr. Rupam Bhattacharyya

Assistant Professor, Dept. of IT

DEPARTMENT OF INFORMATION TECHNOLOGY
GAUHATI UNIVERSITY
GUWAHATI, INDIA
JUNE –2024

GAUHATI UNIVERSITY
DEPARTMENT OF INFORMATION TECHNOLOGY
Gopinath Bordoloi Nagar, Jalukbari Guwahati-781014

DECLARATION

We "***Sanjana Das***", "***Avimanyu Dutta***" Roll No "***200102016***", "***200102019***"a B.Tech. student of the department of Information Technology, Gauhati University hereby declares that I have compiled this report reflecting all my works during the semester long full time project as part of my BTech curriculum.

We declare that we have included the descriptions etc. of my project work, and nothing has been copied/replicated from other's work. The facts, figures, analysis, results, claims etc. depicted in my thesis are all related to my full time project work.

We also declare that the same report or any substantial portion of this report has not been submitted anywhere else as part of any requirements for any degree/diploma etc.

Sanjana Das	Avimanyu Dutta
Branch: IT	Branch: IT
Date: 20-06-2024	Date: 20-06-2024

GAUHATI UNIVERSITY

DEPARTMENT OF INFORMATION TECHNOLOGY

Gopinath Bordoloi Nagar, Jalukbari Guwahati-781014

Date: 20-06-2024

CERTIFICATE

This is to certify that "***Sanjana Das***", "***Avimanyu Dutta***" bearing Roll No: "***200102016***", ***"200102016"*** has carried out the project work "***NetInspect: Exploring CNN based Deep Learning models for Defect Detection from images of PVC specimens and Heat Sink."*** under my supervision and has compiled this report reflecting the candidate's work in the semester long project. The candidate did this project full time during the whole semester under my supervision, and the analysis, results, claims etc. are all related to his/her studies and work during the semester.

I recommend submission of this project report as a part for partial fulfillment of the requirements for the degree of Bachelor of Technology in Information Technology/Computer Science & Engineering of Gauhati University.

Dr. Rupam Bhattacharyya
Assistant Professor
Department of IT, GUIST

GAUHATI UNIVERSITY
DEPARTMENT OF INFORMATION TECHNOLOGY
Gopinath Bordoloi Nagar, Jalukbari Guwahati-781014

External Examiners Certificate

This is to certify that “***Sanjana Das***”, “***Avimanyu Dutta***” bearing Roll No. “200102016”,“200102019” has delivered his project presentation on “20/06/2024” and I examined his/her report entitled “**NetInspect: Exploring CNN based Deep Learning models for Defect Detection from images of PVC specimens and Heat Sink**.” and recommend this project report as a part for partial fulfillment of the requirements for the degree of Bachelor of Technology in Information Technology/Computer Science & Engineering of Gauhati University.

(External Examiner)

GAUHATI UNIVERSITY

DEPARTMENT OF INFORMATION TECHNOLOGY

Gopinath Bordoloi Nagar, Jalukbari Guwahati-781014

TO WHOM IT MAY CONCERN

This is to certify that “**Sanjana Das**”, “**Avimanyu Dutta**” bearing Roll No. “**200102016**”, “**200102019**”, a B.Tech. student of the department of Information Technology, Gauhati University, has submitted the softcopy of his/her project for undergoing screening through anti-plagiarism software and the similar report found to be.

Dr. Rupam Bhattacharyya
Assistant Professor
Department of IT, GUIST

GAUHATI UNIVERSITY
DEPARTMENT OF INFORMATION TECHNOLOGY
Gopinath Bordoloi Nagar, Jalukbari Guwahati-781014

ACKNOWLEDGEMENT

The joy of accomplishment is only complete, when it is accompanied with gratitude to the ones who have helped and encouraged me throughout the process till completion of this project.

We wish to place our deep sense of respect and gratitude to our advisor and Project Guide, **Dr. Rupam Bhattacharyya, Assistant Professor of Department Of Information Technology**, for his motivation and valuable help, which resulted in the successful completion of this project.

We would also like to express our gratitude to **Dr. Shikhar Kumar Sarma**, **Head of the Department of Information Technology** for the smooth conduction and functioning of our classes and projects

We would also like to express our gratitude to the faculty members of the department for their cooperation and generous help for the completion of the project.

Sanjana Das
Branch: IT
Date: 20-06-2024

Avimanyu Dutta
Branch: IT
Date: 20-06-2024

ABSTRACT

Defect detection is crucial in industrial production, significantly impacting both efficiency and product quality. This project explores the efficacy of deep learning models for defect detection on two distinct datasets: the PVC-Infrared dataset for non-destructive evaluation via pulsed thermography, and the Heat Sink Surface Defect Dataset. For the PVC-Infrared dataset, we implement and compare U-Net and SegNet models, while for the Heat Sink Surface Defect dataset, we utilize U-Net, SegNet, and ResNet-50 models. The primary objective is to compare the performance of these models across the different datasets.

The PVC-Infrared dataset, introduced to address the scarcity of publicly available pulsed thermography datasets, benefits from deep learning-based image segmentation to process and analyze the data. Our evaluation focuses on the segmentation accuracy and computational efficiency of U-Net and SegNet models. Conversely, the Heat Sink Surface Defect Dataset is analyzed using U-Net, SegNet, and ResNet-50, focusing on surface defect detection accuracy and processing speed.

Our comparative analysis highlights that preprocessing techniques can significantly enhance model performance by reducing data size without compromising accuracy, especially in the context of pulsed thermography. Additionally, the results demonstrate that the choice of model architecture has a substantial impact on defect detection capabilities, with each model exhibiting varying strengths depending on the dataset characteristics.

Overall, this project underscores the potential of deep learning models to streamline non-destructive evaluation and surface defect detection processes, making them more efficient and accessible to a broader range of industrial applications.

TABLE OF CONTENTS ABSTRACT

LIST OF TABLES:

LIST OF FIGURES:

Chapter 01

INTRODUCTION

1.1 Introduction

In recent years, the integration of artificial intelligence (AI) into industrial applications has significantly enhanced production efficiency and product quality, particularly through the use of defect detection systems. Defect detection plays a crucial role in industrial production, directly affecting both the quality and reliability of the final products. Traditional methods, often reliant on human inspection, are time-consuming and prone to inconsistencies. As a result, automated vision-based defect detection has emerged as a vital area of research and development.

Pulsed thermography (PT) has gained widespread popularity as a non-destructive evaluation (NDE) method due to its high accuracy and speed. It is especially suitable for detecting near-surface defects in materials such as metal, Carbon Fiber Reinforced Polymers (CFRP), and Polyvinyl Chloride (PVC). However, the high complexity of PT data, often containing noise, artifacts, and inconsistencies, presents significant challenges for inspection. Consequently, the application of deep learning methods in processing PT data has become a crucial research topic, aiming to improve the efficiency and accuracy of defect detection. This enhancement is vital for ensuring the reliability and safety of various engineering systems.

In this project, we focus on two specific datasets: the PVC-Infrared dataset and the Heat Sink Surface Defect dataset. The PVC-Infrared dataset addresses the need for improved defect detection in non-destructive evaluation via pulsed thermography. To analyze this dataset, we implement and compare the performance of U-Net and SegNet models. These models are evaluated based on their ability to segment and interpret complex PT data, thereby facilitating more accurate defect detection.

The Heat Sink Surface Defect dataset pertains to the detection of surface defects in industrial components, such as tungsten copper alloy heat sinks used in chip heat dissipation. Ensuring the integrity of the heat sink's surface coating is critical, as defects can lead to thermal failure of the chip. Traditional manual inspection methods are labor-intensive and dependent on the proficiency of the inspectors. Therefore, automated inspection using deep learning models offers a more efficient and reliable solution. For this dataset, we employ U-Net, SegNet, and ResNet-50 models, evaluating their performance in terms of detection accuracy and processing speed.

The primary motivation behind this project is to explore and compare the performance of different deep learning models across varied datasets, highlighting the impact of preprocessing techniques and model architecture on defect detection capabilities. By analyzing the results, we aim to identify the most efficient and accurate methods for defect detection in different contexts, thereby contributing to the optimization of industrial inspection processes.

This project not only advances the state of the art in defect detection but also provides valuable insights into the application of deep learning in complex, real-world scenarios. By demonstrating the potential of deep learning models to streamline these processes, we aim to make them more accessible and efficient for a wider range of industrial applications. The outcomes of this research have practical implications for improving the quality and reliability of industrial production, ultimately contributing to advancements in non-destructive evaluation and automated defect detection.

1.2. Objective

The objective of this project is to evaluate and compare the performance of deep learning models in defect detection using the PVC-Infrared dataset and the Heat Sink Surface Defect dataset. Specifically, the project aims to:

1. **Implement and Evaluate Models on PVC-Infrared Dataset**:
   - Utilize U-Net and SegNet models to analyze and segment pulsed thermography data.
   - Assess the models' accuracy and efficiency in detecting near-surface defects in materials such as PVC.
2. **Implement and Evaluate Models on Heat Sink Surface Defect Dataset**:
   - Apply U-Net, SegNet, and ResNet-50 models to detect surface defects in tungsten copper alloy heat sinks.
   - Compare the models' performance in terms of detection accuracy and processing speed.
3. **Compare Performance Across Models and Datasets**:
   - Conduct a comparative analysis to identify the most effective deep learning models for each dataset.
   - Evaluate the impact of preprocessing techniques on model performance, aiming to optimize data handling and reduce computational load.
4. **Enhance Defect Detection Processes**:
   - Improve the accuracy and reliability of defect detection in non-destructive evaluation and industrial inspection.
   - Provide insights into the application of deep learning in real-world industrial scenarios, contributing to the development of more efficient automated inspection systems.

By achieving these objectives, the project seeks to advance the field of defect detection, making it more accessible and efficient for various industrial applications, thereby enhancing overall production quality and reliability.

1.3. Background:

1.3.1 Non-Destructive Evaluation (NDE)

Non-Destructive Evaluation (NDE) is a crucial field in engineering that involves assessing materials, components, or structures without causing damage. NDE techniques are employed to detect and evaluate defects or irregularities, ensuring the safety and reliability of various systems and structures. Pulsed thermography (PT) is one such NDE method known for its non-invasive nature and high accuracy, making it particularly suitable for defect detection in materials like metal, Carbon Fiber Reinforced Polymers (CFRP), and Polyvinyl Chloride (PVC).

1.3.2 Surface Defect Detection

Surface defect detection plays a vital role in industrial production, directly affecting the quality and reliability of products. Traditional methods of surface defect detection often rely on manual inspection, which is time-consuming and inconsistent. Automated vision-based defect detection systems have emerged as essential tools for improving inspection efficiency and accuracy. These systems leverage advanced image processing and deep learning techniques to detect surface defects in various industrial components.

1.3.3 Characteristics of Image Segmentation

Image segmentation is a fundamental task in computer vision that involves dividing an image into meaningful and distinct regions. In the context of defect detection, accurate segmentation is vital for precisely identifying the location and extent of defects within an image. Key characteristics of image segmentation include precise localization, the ability to handle complex and varied image data, and the capability to process large volumes of data efficiently.

1.3.4 CNN-Based Image Segmentation

Convolutional Neural Networks (CNNs) have demonstrated exceptional capabilities in image processing tasks, including segmentation. CNN-based image segmentation, such as the U-Net, SegNet, and ResNet-50 architectures, is particularly effective in tasks that require detailed localization, making them well-suited for defect detection. The hierarchical and convolutional nature of CNNs enables them to learn intricate patterns and features within images, enhancing the accuracy of segmentation tasks.

1.3.5 PVC-Infrared Dataset Description

The PVC-Infrared dataset is designed for non-destructive evaluation via pulsed thermography. It provides infrared images of PVC materials with annotated defect regions. This dataset addresses the challenges associated with the high complexity and noise present in PT data, offering a foundation for training and evaluating deep learning models for defect detection. The dataset's availability facilitates research and development in applying deep learning techniques to NDE data.

1.3.6 Heat Sink Surface Defect Dataset Description

The Heat Sink Surface Defect dataset pertains to the detection of surface defects in tungsten copper alloy heat sinks used in chip heat dissipation. Ensuring the integrity of the heat sink’s surface coating is critical, as defects can lead to thermal failure of the chip. This dataset provides annotated images of heat sink surfaces, serving as a benchmark for evaluating deep learning models' performance in surface defect detection tasks.

Chapter 02

LITERATURE REVIEW

2.1 Introduction

This chapter aims to analyze the existing literature on defect detection using deep learning techniques applied to pulsed thermography and surface imaging datasets. The goal is to highlight the progress made, identify common themes, and suggest directions for future research. Defect detection is a crucial aspect of quality control in various industries, and advancements in deep learning have significantly enhanced the accuracy and efficiency of these processes. This review will summarize key findings from the literature and underscore the importance of employing advanced models such as U-Net, SegNet, and ResNet-50 for defect detection in PVC-Infrared and Heat Sink Surface Defect datasets.

2.2 Significance of Defect Detection

Defect detection is vital for ensuring product quality and reliability across multiple industries, including aerospace, automotive, and electronics. Detecting and addressing defects early in the production process can prevent costly failures and enhance safety. For instance, in the aerospace industry, detecting defects in composite materials is crucial for maintaining structural integrity. In electronics, identifying surface defects in heat sinks is essential to prevent thermal failures in chips.

2.3 Pulsed Thermography and Surface Imaging

Pulsed thermography (PT) and surface imaging are prominent non-destructive evaluation (NDE) methods. PT uses thermal waves to detect subsurface defects, making it highly effective for materials like PVC, CFRP, and metals. Surface imaging, especially for materials like tungsten copper alloy used in heat sinks, relies on capturing detailed surface features to identify defects. These methods provide rich datasets that, when combined with deep learning, can significantly improve defect detection accuracy.

2.4 Previous Approaches in Defect Detection

Traditional defect detection methods include manual inspection and simple thresholding techniques, which are time-consuming and prone to human error. Advanced methods have employed computer vision and machine learning to improve detection accuracy. For example, statistical detection and wavelet transform-based methods have been used in industrial environments. However, these techniques often require manual parameter tuning and lack the robustness of modern deep learning approaches.

2.5 Role of Deep Learning in Defect Detection

Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized defect detection by automating feature extraction and improving segmentation accuracy. Key architectures like U-Net, SegNet, and ResNet-50 have demonstrated exceptional performance in various image processing tasks. These models are capable of handling complex data and providing precise localization and classification of defects.

- **U-Net**: Known for its symmetric encoder-decoder structure, U-Net excels in image segmentation tasks that require detailed localization. It has been widely used in medical imaging and industrial defect detection.
- **SegNet**: Optimized for pixel-wise image segmentation, SegNet is computationally efficient and effective in defect detection applications.
- **ResNet-50** : Leveraging its deep residual learning framework, ResNet-50 serves as a robust feature extractor tailored for high-accuracy segmentation tasks.

2.6 Challenges in Deep Learning for Defect Detection

Several challenges persist in applying deep learning to defect detection:

- **Data Quality and Availability**: High-quality annotated datasets are essential but often limited in specialized fields like NDE. Imbalanced datasets, where defective samples are rare, pose significant challenges.
- **Data Complexity**: PT data often contains noise and artifacts, requiring robust preprocessing techniques.

- **Model Generalization**: Ensuring models generalize well across different defect types and materials is crucial for practical applications.

2.7 Applications and Comparative Analysis

Deep learning models have been applied to various defect detection tasks with notable success:

- **PVC-Infrared Dataset**: This dataset, capturing thermal images of PVC materials, has been used to train and evaluate models like U-Net and SegNet.
- **Heat Sink Surface Defect Dataset**: Models such as U-Net, SegNet, and ResNet-50 serving as an encoder backbone, have been applied to this dataset to detect surface defects in tungsten copper alloy heat sinks.

Comparative studies have shown that appropriate preprocessing techniques and model selection can significantly enhance detection performance. For instance, U-Net and SegNet have demonstrated high accuracy in segmenting defect regions in infrared images, while ResNet-50 has been effective in classifying surface defects.

2.8 Summary of Findings

The literature review underscores the importance of advanced imaging techniques and deep learning models in defect detection. Integrating PT and surface imaging with deep learning offers significant potential for improving accuracy and efficiency. The comparative analysis of models like U-Net, SegNet, and ResNet-50 (serving as encoder backbone) across different datasets highlights their strengths and areas for improvement. This project aims to contribute valuable insights and advancements to the field, paving the way for more reliable and efficient industrial inspection processes.

Table 1: Summary of Literature Review

Model	Dataset	Metric considered for defect detection	Reference and remarks
U-Net	PVC Infrared	Average IoU	[3] and state of the architecture performance over PVC
SegNet	PVC Infrared	Average IoU	[3] and state of the architecture performance over PVC
U-Net VGG11	PVC Infrared	Average IoU	[3] and state of the architecture performance over PVC
SegNet VGG11	PVC Infrared	Average IoU	[3] and state of the architecture performance over PVC
U-Net	Heat Sink	Accuracy, Precision, Recall, IoU, F1-score, FLOPs	[4] and state of the architecture performance over Heat Sink.
GSLU-Net	Heat Sink	Accuracy, Precision, Recall, IoU, F1-score, FLOPs	[4] and state of the architecture performance over Heat Sink.
SE light U-Net	Heat Sink	Accuracy, Precision, Recall, IoU, F1-score, FLOPs	[4] and state of the architecture performance over Heat Sink.

Chapter 03

PROBLEM STATEMENT AND CHALLENGES

3.1. Problem Statement

NetInspect: Exploring CNN Architectures for Defect Detection in Diverse Materials

3.2. Objectives:

The objective of this project is to evaluate and compare the performance of deep learning models in defect detection using the PVC-Infrared dataset and the Heat Sink Surface Defect dataset. Specifically, the project aims to:

1. **Implement and Evaluate Models on PVC-Infrared Dataset**:
   - Utilize U-Net and SegNet models to analyze and segment pulsed thermography data containing PVC specimens.
   - Assess the models' accuracy and efficiency in detecting near-surface defects in materials such as PVC.
2. **Implement and Evaluate Models on Heat Sink Surface Defect Dataset**:
   - Apply U-Net, SegNet, and ResNet-50 models to detect surface defects in tungsten copper alloy heat sinks.
   - Compare the models' performance in terms of detection accuracy.
3. **Compare Performance Across Models and Datasets**:
   - Conduct a comparative analysis to identify the most effective deep learning models for each dataset.
   - Evaluate the impact of preprocessing techniques on model performance, aiming to optimize data handling and reduce computational load.

4. **Enhance Defect Detection Processes**:
    - Improve the accuracy and reliability of defect detection in non-destructive evaluation and industrial inspection.
    - Provide insights into the application of deep learning in real-world industrial scenarios, contributing to the development of more efficient automated inspection systems.

By achieving these objectives, the project seeks to advance the field of defect detection, making it more accessible and efficient for various industrial applications, thereby enhancing overall production quality and reliability.

3.3. Challenges:

- **Limited Computing Power**: Not enough computational resources slowed down model training, requiring exploration of alternative options.
- **Complex Infrared Dataset Handling**: Understanding thermal imaging data was tough, needing special preprocessing for effective defect detection.
- **Model Complexity**: Grasping different models like U-Net, SegNet, and ResNet-50 required extensive study.
- **Training Time Constraints**: Lengthy training sessions restricted exploring different model configurations and hyperparameters.
- **Collaboration Challenges**: Coordinating tasks and maintaining effective communication within a multidisciplinary team was challenging.
- **Generalization Across Datasets**: Ensuring models worked well on diverse datasets with various defect types, sizes, and orientations required careful consideration and strategy adjustments.
- **Dataset Size and Quality Constraints**: The PVC dataset's small size and reliance on infrared images posed challenges, needing meticulous preprocessing and model adaptation.
- **Model Comparison Process**: Comparing performance among different models required rigorous experimentation and evaluation protocols.

- **Hyperparameter Tuning Complexity**: Finding the optimal hyperparameters for each model involved extensive trial and error due to the complex nature of the datasets and tasks.

Chapter 04

METHODOLOGY

4.1 Overview of our methodology

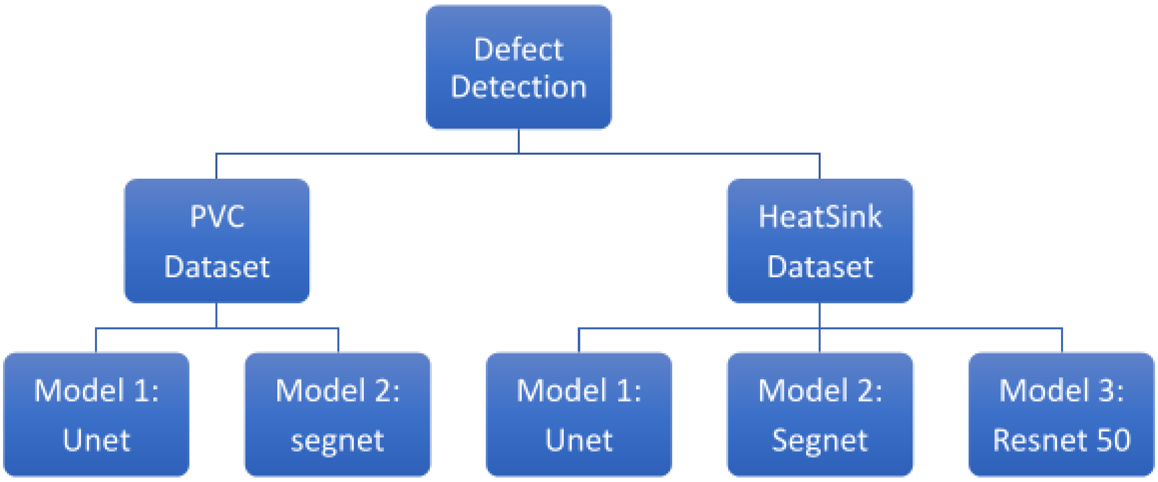

Figure 1: Hierarchical organisation of our methodology

Figure 2: System pipeline

4.2 Specimens' Information and Data Acquisition

4.2.1 PVC-Infrared Dataset In this study, we introduce the PVC-Infrared dataset, a deep learning dataset consisting of 19 thermal image sequences for 19 PVC specimens. Each specimen has dimensions of 100 mm × 100 mm × 5 mm, with cylindrical holes of varying sizes and depths created on the bottom side to simulate subsurface defects. The diameter of the cylindrical holes ranges from 2 mm to 10 mm, and the depth ranges from 2.5 mm to 4.5 mm from the bottom surface. Figure 1 provides CAD images of eight specimens viewed from the bottom.

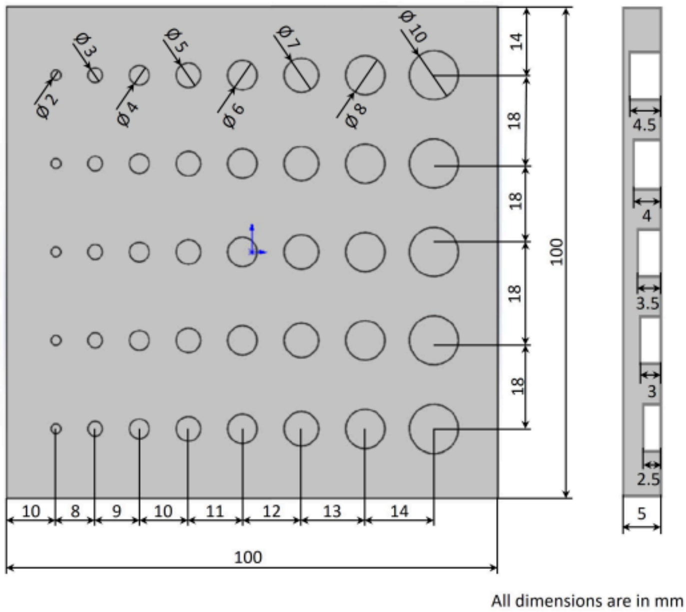

Figure 3: An example of specimens with defects in all possible sizes in all possible locations

4.2.2 Heat Sink Surface Defect Dataset We have also utilized the Heat Sink Surface Defect dataset, containing 1000 surface images of gold-plated tungsten copper alloy heat sinks with defects. Each image contains multiple defects and is annotated at the pixel level, providing detailed information about the surface imperfections essential for training and validating defect detection models.

4.3 Data Preprocessing

4.3.1 PVC-Infrared Dataset Preprocessing Data preprocessing is a pivotal phase in readying the dataset for model training. For the PVC-Infrared dataset, we initialize a custom Dataset class tailored for loading infrared images and their corresponding masks. This class incorporates essential functionalities such as image normalization, transformation, and efficient data loading. Specific steps include:

- **Normalization**: Scaling pixel values to a standard range to improve model convergence.
- **Augmentation**: Applying random transformations such as rotations, flips, and zooms to enhance the robustness of the model.
- **Mask Preparation**: Ensuring defect masks are correctly aligned and formatted for segmentation tasks.

4.3.2 Heat Sink Surface Defect Dataset Preprocessing For the Heat Sink Surface Defect dataset, preprocessing involves the following steps:

1. **Contrast Enhancement**: Enhancing the contrast of defect masks using histogram equalization for better readability and detection.
2. **Normalization**: Scaling image pixel values to a range of [0, 1].
3. **Binarization**: Ensuring masks are binary by setting a threshold to distinguish defect regions from the background.
4. **Splitting Data**: Dividing the dataset into training and validation sets using an 80-20 split for effective model evaluation.

4.4 Model Architecture Overview

4.4.1 U-Net Architecture

U-Net is a popular and effective neural network approach for biomedical image segmentation. It is composed of two key components: the encoding stage and the decoding stage. The encoding stage involves several consecutive blocks of convolutional neural network operations, each followed by a max pooling operation. The decoding stage consists of an equal number of transposed convolution blocks to upsample the output back to the same dimensions as the input data. Importantly, U-Net also includes skip connections, which facilitate the transfer of high-resolution information from the encoding stage to the decoding stage by performing copy and concatenation of feature maps. In this work, we employ a lightweight version of U-Net, with a single convolutional layer used as an operation block, as shown in Figure

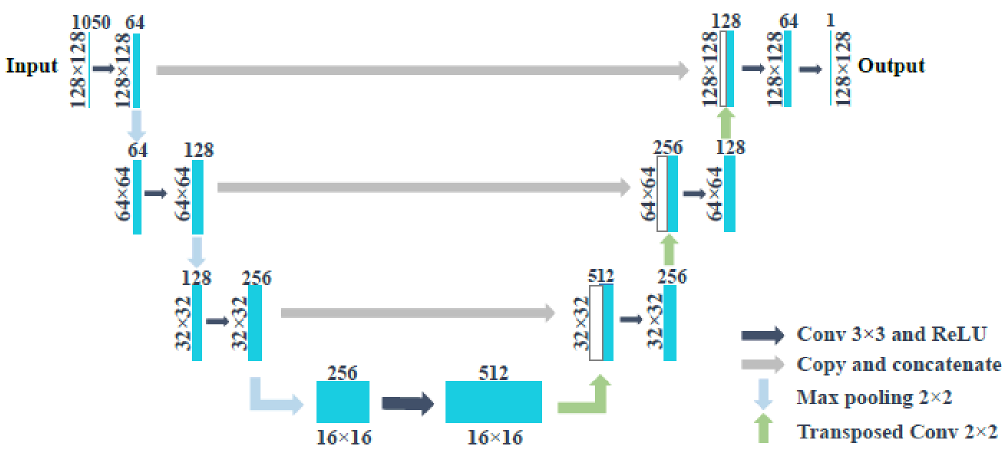

Figure 4. U-Net architecture presented in [3]

4.4.2 SegNet Architecture

SegNet, is similar to the U-Net architecture, which has encoding and decoding stages. The main difference is that, in SegNet, only the locations where max pooling is performed, which can also be referred to as pooling indices, are copied and used for upsampling to the decoding stage at the skip connection instead of copying the entire feature map from the encoding stage. Compared to the original SegNet, we have employed fewer steps both at the encoding and decoding stages. The model architecture used is illustrated in Figure 7.

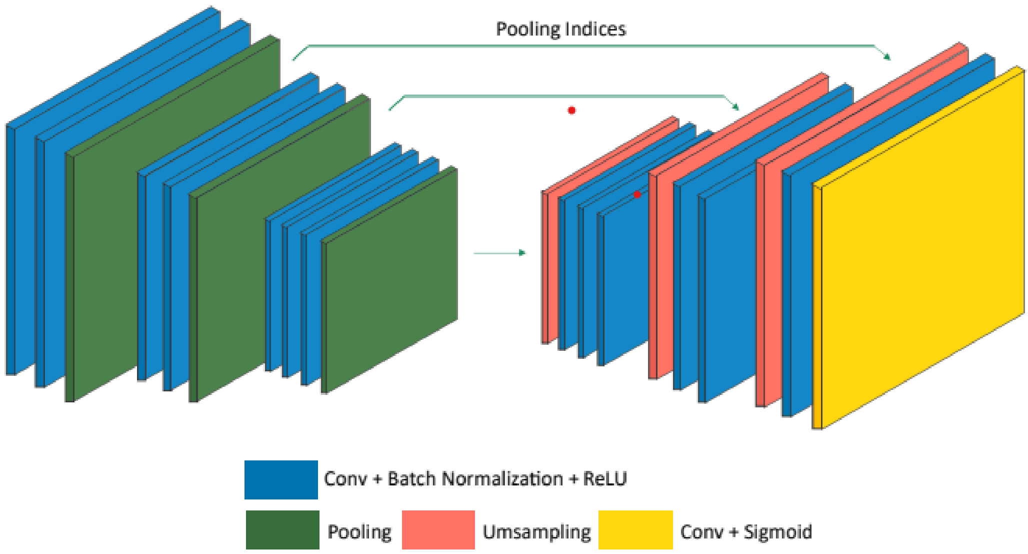

Figure 5. SegNet architecture adapted from a figure in [3]

4.4.3 Resnet 50 Backbone Architecture

ResNet-50 is a robust deep learning architecture that employs Residual Learning, essential for training very deep networks effectively. The "50" denotes the number of layers in the network. Deep convolutional networks have achieved significant advancements in image classification by increasing the layer count to tackle complex tasks and enhance accuracy. However, as networks deepen, accuracy often plateaus and can even decline, a problem that Residual Learning mitigates. Instead of learning features directly, Residual Learning targets the residuals—the differences between the learned features and the input to that layer. ResNet-50 accomplishes this via shortcut connections that directly link the input of a layer to a subsequent layer, easing the training of deep networks and overcoming the issue of diminishing accuracy. This approach, utilizing residuals and shortcut connections strategically, enables ResNet-50 to sustain high accuracy and performance even as the network depth grows.

In the context of image segmentation, ResNet-50 serves as the encoder in an encoder-decoder architecture, effectively extracting high-level features from the input image. The encoder downsamples the input through a series of convolutional and pooling layers, generating a low-resolution feature map. The decoder then up samples this feature map back to the original image size using a series of upsampling and convolutional layers, producing a detailed segmentation mask. This method leverages the powerful feature extraction capabilities of ResNet-50 while addressing the need for precise pixel-wise classification, making it well-suited for segmentation tasks.

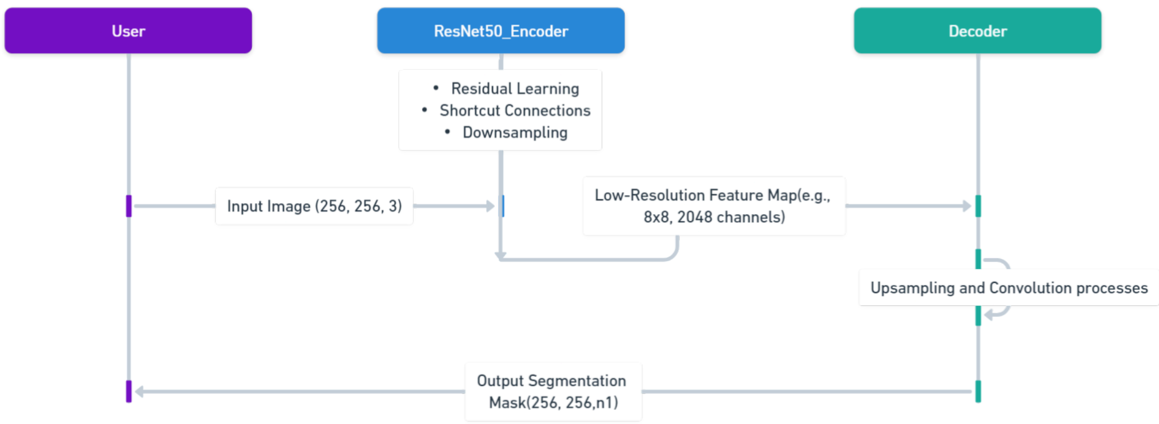

Figure 6: Resnet 50 backbone segmentation architecture

4.5 Training Process

4.5.2 PVC-Infrared Dataset Training

For the PVC-Infrared dataset, the training process involved feeding the preprocessed data into the U-Net and SegNet models to learn defect patterns. This process included optimizing the model parameters using suitable loss functions and employing backpropagation to update the network weights. The primary goal was to minimize the difference between the predicted and actual defect masks, thereby enhancing the models' ability to accurately identify and localize defects.

4.5.2 Heat Sink Surface Defect Dataset Training

For the Heat Sink Surface Defect dataset, the training process involved using the U-Net, SegNet, and ResNet-50 models. Similar to the PVC-Infrared dataset training, the models were fed preprocessed data to learn defect patterns. The optimization of model parameters, using appropriate loss functions and backpropagation, aimed to reduce the discrepancy between predicted and actual defect masks, ensuring the models effectively identify and localize defects.

4.6 Validation Process

4.6.1 PVC-Infrared Dataset Validation

During the training of the PVC-Infrared dataset, a separate validation set was used to assess the performance of the U-Net and SegNet models on unseen data. This step was crucial for preventing overfitting and ensuring the models could generalize well to new data. Validation metrics such as loss and accuracy were monitored to guide the training process and fine-tune the models.

4.6.2 Heat Sink Surface Defect Dataset Validation

For the Heat Sink Surface Defect dataset, the U-Net, SegNet, and ResNet-50 models were validated using a separate set of data. This validation process involved monitoring various performance metrics to ensure the models could generalize effectively and to adjust hyperparameters as needed, thus enhancing model robustness.

4.7 Testing the Model

4.7.1 PVC-Infrared Dataset Testing

After training, the U-Net and SegNet models were tested on a separate set of images not used during training or validation. The performance was evaluated using the same metrics employed during validation to provide insights into the models' abilities to generalize to unseen data.

4.7.2 Heat Sink Surface Defect Dataset Testing

Similarly, the U-Net, SegNet, and ResNet-50 models were tested on a distinct set of unseen images from the Heat Sink Surface Defect dataset. Various metrics were used to assess the models' performances and to evaluate their generalization capabilities.

48 Evaluation Metrics

To comprehensively assess the performance of the models across both datasets, a set of evaluation metrics was employed:

- **IoU (Intersection over Union)**: Measures the overlap between predicted and actual defect regions.
- **Dice Coefficient**: Quantifies the similarity between predicted and ground truth masks.
- **Pixel Accuracy**: Evaluates the percentage of correctly classified pixels.
- **Precision, Recall, and F1-score**: Used for classification tasks to provide a balanced view of model performance.

The experimental setup, including the specific parameters such as learning rate, batch size, and dropout rates, were crucial elements in achieving the reported results. This transparency in the experimental setup enhances the reproducibility and trustworthiness of the study.

Chapter 05

EXPERIMENTATION

5.1 Brief Overview

The implementation of this project involves several key steps, from data acquisition and preprocessing to model training and evaluation, for two different datasets: PVC-Infrared and Heat Sink Surface Defect. For the PVC-Infrared dataset, U-Net and SegNet models are implemented, while for the Heat Sink Surface Defect dataset, U-Net, SegNet, and ResNet-50 models are employed. This chapter details the processes and methodologies for each dataset.

5.2 Experimental Setup

5.2.1 Hyperparameters

- **Learning Rate**: Set at 0.001 for optimal convergence.
- **Batch Size**: A batch size of 16 is used to balance memory usage and training speed.

5.2.2 Training Environment

- **Hardware**: The training is conducted on GPUs to accelerate computation.
- **Software**: TensorFlow and Keras are used for model implementation, leveraging their efficient handling of neural network operations.

5.3 Evaluation Metrics:

We utilize accuracy, precision, recall, intersection-over-union (IoU), and F1-score as key metrics to evaluate the performance of our defect segmentation method. These metrics are defined as follows:

- Precision measures the proportion of predicted defects that are actually true defects.
- Recall measures the proportion of actual defects that are correctly identified by the model.
- F1-score is a harmonic mean of precision and recall, providing a balanced view of model performance.
- Accuracy represents the overall percentage of correctly classified pixels (defect and non-defect).

$$\text{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$
$$\text{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$
$$\text{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$
$$\text{IoU} = \frac{\text{target} \cap \text{prediction}}{\text{target} \cup \text{prediction}} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN} + \mathrm{FP}}$$
$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

The segmentation results are categorized into:

1. True positive (TP): accurately segmented defect pixels.
2. False positive (FP): incorrectly segmented non-defect pixels as defect pixels.
3. False negative (FN): defect pixels that were missed in the segmentation.

In defect detection for industrial products, high precision indicates a low false alarm rate for defect segmentation, while high recall denotes a low rate of missed defect detection. The F1-score provides a balanced assessment of precision and recall, essential for evaluating the effectiveness of our defect segmentation method. IoU is particularly critical for evaluating the spatial accuracy and delineation of defect regions by our segmentation model, especially relevant for applications in industrial product quality control.

 Model Building:

5.4 PVC-Infrared Dataset

5.4.1 Modules and Libraries Used:

- **NumPy (np):** Provides fundamental numerical operations on multidimensional arrays for efficient data manipulation and calculations.
- **Pandas (pd):** Offers powerful data structures (Series and DataFrames) for data manipulation, analysis, and visualization, facilitating tasks like loading, cleaning, and exploring datasets.
- **Matplotlib (plt):** A versatile library for creating various static, animated, and interactive visualizations, including plots, histograms, and charts, to gain insights from the data.
- **Scikit-learn (sklearn):** A collection of machine learning algorithms for tasks like classification, regression, clustering, and model selection. While not directly used for deep learning model building, it can be helpful for data preprocessing or preparing baselines to compare against neural network performance.
- **OS (os):** Provides functionalities for interacting with the operating system, including file system navigation, directory operations, and process management. This can be useful for loading data from files or directories.

5.4.2 Data Acquisition and Preprocessing

- **Dataset Description**: The PVC-Infrared dataset consists of 19 thermal image sequences for 19 PVC specimens. Each specimen has dimensions of 100 mm × 100 mm × 5 mm, with cylindrical holes of varying sizes and depths to simulate subsurface defects.
- **Data Loading**: Images and masks are loaded using a custom Dataset class to ensure efficient handling and processing.

- **Normalization**: Image pixel values are scaled to a standard range to improve model training stability.
- **Augmentation**: Random transformations, such as rotations and flips, are applied to the images to enhance the model's ability to generalize.

5.4.3 Model Architecture

5.4.3.1 U-Net Architecture

The U-Net architecture for the PVC dataset segmentation consists of an encoder-decoder network designed to capture and reconstruct defect patterns from infrared images. Here's a detailed breakdown based on the provided code:

- **Encoder:**
  - The encoder stage begins with a series of convolutional layers (one_step_conv) that progressively extract features from the input image (x). Each convolutional layer is followed by batch normalization and ReLU activation, enhancing the network's ability to learn complex features.
  - There are four down-sampling stages (conv_down_1 to conv_down_4) using max-pooling layers (pool1 to pool4). These stages reduce the spatial dimensions of the feature maps while increasing the number of feature channels, allowing the network to extract hierarchical representations of the input image.
- **Decoder:**
  - The decoder stage aims to upsample the feature maps back to the original input dimensions while preserving the spatial information necessary for precise segmentation.
  - Up-sampling is performed using transposed convolutional layers (nn.ConvTranspose2d). These layers increase the spatial resolution of the feature maps and reduce the number of channels to match the target output.
  - At each up-sampling stage (upsample_1 to upsample_4), the up-sampled feature maps are concatenated (torch.cat) with the corresponding feature

maps from the encoder stage (down_4 to down_1). This process creates skip connections, allowing the network to retain fine-grained details and spatial context from the encoding stage.
    - Each concatenated feature map undergoes another set of convolutional operations (conv_up_1 to conv_up_4) to further refine the segmented output.

- **Output Layer:**
    - Finally, the output of the last convolutional layer (conv_out) is passed through a sigmoid activation function (nn.Sigmoid()) to produce a probability map where each pixel indicates the likelihood of belonging to a defect region.
- **Skip Connections:**
    - Skip connections play a crucial role in U-Net by facilitating the flow of high-resolution information from the encoder to the decoder stages. This helps in precise localization of defects by enabling the network to merge both low-level and high-level features effectively.

This architecture leverages the strengths of both convolutional and transposed convolution operations, combined with skip connections, to achieve accurate segmentation of defects in the PVC-Infrared dataset. The design ensures that the U-Net model can effectively capture and reconstruct intricate defect patterns while maintaining spatial context and minimizing information loss throughout the network.

5.4.3.2 SegNet Architecture

SegNet, a convolutional neural network architecture designed for semantic segmentation tasks, consists of two main components: the encoder and the decoder.

- **Encoder**: Similar to U-Net, SegNet's encoder comprises multiple convolutional layers followed by max-pooling operations. This sequence extracts hierarchical feature maps that capture different levels of abstraction from the input data. In the provided code, the encoder begins with convolutional layers of increasing depth (64 to 512 channels), each followed by max-pooling layers (2x2) to downsample the spatial dimensions.

- **Decoder**: In SegNet, the decoder upsamples the feature maps to the original input dimensions using pooling indices stored during the max-pooling operation in the encoder phase. This step aids in precise boundary recovery by maintaining spatial information. The decoder section in the code utilizes transposed convolutional layers (Conv2DTranspose) to upsample the feature maps. Each upsample operation is followed by additional convolutional layers to refine the feature representation. This process gradually recovers the spatial resolution lost during the encoding stage.
- **Output Layer**: The final layer of the SegNet architecture is a convolutional layer with a sigmoid activation function. This layer produces pixel-wise class probabilities, making it suitable for binary segmentation tasks. In the provided code, the output layer has a single channel (1x1 convolution) with a sigmoid activation, generating an output mask where each pixel indicates the probability of belonging to the defect class.

SegNet's architecture is specifically tailored for efficient and effective semantic segmentation, leveraging both encoding for feature extraction and decoding for precise spatial localization of defects within the PVC-Infrared dataset. This structured approach ensures that the model can accurately identify and delineate defect regions in the infrared images, essential for quality control and defect detection applications.

5.4.4 Training Process

- **Training Setup**: Both U-Net and SegNet models are trained using the preprocessed PVC-Infrared dataset.
- **Optimization**: The Adam optimizer with a learning rate of 0.001 is used, and the models are trained for 50 epochs with a batch size of 16.
- **Callbacks**: Early stopping and model checkpoint callbacks are employed to prevent overfitting and save the best model.

5.4.5 Validation Process

- **Validation Metrics**: During training, validation loss and Intersection over Union (IoU) are monitored to assess model performance on unseen data.

- **Hyperparameter Tuning**: The validation results guide the tuning of hyperparameters to improve model performance.

5.4.6 Testing the Models

- **Test Setup**: The trained U-Net and SegNet models are tested on a separate set of images not used during training or validation.
- **Evaluation Metrics**: Performance is evaluated using metrics such as IoU, Dice coefficient, and pixel accuracy.

5.5 Heat Sink Surface Defect Dataset

5.5.1 Modules and Libraries Used:

- **NumPy (np):** Provides fundamental numerical operations on multidimensional arrays for efficient data manipulation and calculations.
- **Matplotlib (plt):** A versatile library for creating various static, animated, and interactive visualizations, including plots, histograms, and charts, to gain insights from the data.
- **TensorFlow (tf):** A comprehensive deep learning framework for creating, training, and deploying a wide range of neural network architectures. It provides building blocks (layers, optimizers, losses, etc.) and APIs for efficient neural network development.
- **Keras (tf.keras):** A high-level API built on top of TensorFlow, offering a user-friendly interface for building and training neural networks. It simplifies the process by providing pre-built layers, optimizers, losses, and metrics, making it easier to get started with deep learning.
- **OpenCV (cv2):** A real-time computer vision library for image and video processing, providing algorithms and functions for tasks like object detection, image manipulation, and feature extraction.
- **Scikit-learn (sklearn):** A collection of machine learning algorithms for tasks like classification, regression, clustering, and model selection. While not directly

used for deep learning model building, it can be helpful for data preprocessing or preparing baselines to compare against neural network performance.

- **OS (os):** Provides functionalities for interacting with the operating system, including file system navigation, directory operations, and process management. This can be useful for loading data from files or directories.

5.5.2 Data Acquisition and Preprocessing

- **Dataset Description**: This dataset includes 1000 surface images of gold-plated tungsten copper alloy heat sinks, each annotated at the pixel level to highlight multiple defects.
- **Data Loading and Contrast Enhancement**: Images and masks are loaded, and contrast enhancement is applied to the masks for better readability and segmentation accuracy.
- **Normalization and Binarization**: Images are normalized to the range [0, 1], and masks are binarized to ensure clear distinction between defect and non-defect regions.
- **Data Splitting**: The dataset is split into training and validation sets using an 80-20 split to facilitate effective model training and evaluation.

5.5.3 Model Architecture

5.5.3.1 U-Net Architecture

the U-Net model implemented for segmenting heatsink defects in the provided Heatsink dataset. The description focuses on the model architecture, addressing the additional layers, loss function, and their functionalities.

The implemented U-Net follows the classic encoder-decoder structure:

- **Encoder:** This part captures features from the input image. It consists of five convolutional blocks with the following steps in each block:

- Two consecutive 3x3 convolutional layers with ReLU activation for feature extraction.
- A 2x2 max-pooling layer for downsampling the feature maps, reducing their spatial dimensions while increasing the number of channels.

- **Decoder:** This part upsamples the captured features and refines them to produce the segmentation mask. It consists of five upsampling blocks with the following steps in each block:
    - Upsampling the feature map using a 2x2 UpSampling2D layer, effectively doubling its spatial dimensions.
    - Concatenation with the corresponding feature map from the encoder path at the same spatial resolution. This helps retain spatial information lost during downsampling.
    - Two consecutive 3x3 convolutional layers with ReLU activation for refining the combined features.

Additional Layers:

- **Dropout:** A dropout layer with a rate of 0.5 is added after the final convolutional block in the encoder. This helps prevent overfitting by randomly dropping a certain percentage of activations during training, forcing the model to learn robust features that are not dependent on specific neurons.
- **Batch Normalization:** Batch normalization layers are not explicitly shown in the provided code, but they are commonly used after each convolutional layer in U-Net architectures. These layers normalize the activations of the previous layer, improving training stability and gradient flow.

Loss Function:

The model employs the Dice coefficient loss function for optimization. The Dice coefficient measures the overlap between the predicted and actual defect regions. It is particularly well-suited for segmentation tasks as it penalizes models for both missing true defects (false negatives) and predicting defects where there are none (false positives). The implemented loss function calculates 1 minus the Dice coefficient, effectively minimizing the difference between the predicted and actual segmentation masks.

5.5.3.2 SegNet Architecture

the SegNet model implemented for segmenting heatsink defects in the Heatsink dataset. The description focuses on the model architecture, highlighting the encoder, decoder, output layer, and their functionalities.

SegNet follows an encoder-decoder structure similar to U-Net, but with key differences in the decoder pathway:

- **Encoder:**
    - o Similar to U-Net, the encoder consists of convolutional blocks for feature extraction.
    - o Each block uses two consecutive 3x3 convolutional layers with ReLU activation for feature extraction followed by Batch Normalization for improved training stability.
    - o Max pooling layers are used for downsampling the feature maps, reducing their spatial dimensions while increasing the number of channels.
- **Decoder:**
    - o Unlike U-Net, SegNet utilizes upsampling layers to increase the spatial resolution of the feature maps.
    - o However, to recover precise spatial boundaries during upsampling, SegNet incorporates **pooling indices** from the corresponding encoder block. These indices essentially store the locations of maximum activations in the pooling operation, allowing for more accurate localization of features in the upsampled decoder pathway.
    - o Two consecutive 3x3 convolutional layers with ReLU activation and Batch Normalization are applied after each upsampling step to refine the upsampled features.
- **Output Layer:**

    - A final 1x1 convolutional layer with sigmoid activation is used to generate pixel-wise class probabilities for each pixel in the image. This is suitable for segmentation tasks, as the output represents the likelihood of each pixel belonging to the defect class.

Overall, the SegNet architecture leverages pooling indices from the encoder to achieve precise upsampling in the decoder, making it effective for segmenting objects with intricate boundaries.

5.5.3.3 ResNet-50 Backbone Architecture

the ResNet-50 based segmentation model implemented for segmenting heatsink defects in the Heatsink dataset. The description focuses on the model architecture, highlighting the pre-trained feature extractor (ResNet-50) and the custom decoder network.

The model leverages transfer learning by utilizing a pre-trained ResNet-50 model for feature extraction and a custom decoder network for upsampling and segmentation:

- **Feature Extraction:**
    - The core of the model is a pre-trained ResNet-50 convolutional neural network (CNN) without the final classification layers (include_top=False).
    - ResNet-50 is a deep learning architecture known for its residual learning capabilities, which help to address the vanishing gradient problem and enable effective training of deeper networks.
    - During training, the weights of the pre-trained ResNet-50 are frozen, essentially using it as a feature extractor that captures generic image features from the input heatsink images.
- **Decoder:**
    - A custom decoder network is appended to the output of the pre-trained ResNet-50.
    - The decoder employs several UpSampling2D layers to progressively increase the spatial resolution of the extracted features. This is crucial as

    - the features extracted by ResNet-50 have a lower resolution compared to the original input image.
    - Following each upsampling step, a convolutional layer with ReLU activation is applied to refine the upsampled features.
    - The number of filters in the convolutional layers progressively decreases as the decoder moves up, typically corresponding to higher-level semantic information.
- **Output Layer:**
    - A final 1x1 convolutional layer with sigmoid activation is used to generate a binary segmentation map. The sigmoid function outputs values between 0 and 1, representing the probability of each pixel belonging to the defect class.

Overall, the ResNet-50 segmentation model exploits the power of pre-trained features from ResNet-50 and utilizes a custom decoder network to upsample and classify pixels for heatsink defect segmentation.

5.5.4 Training Process

Training Setup:

- All three models (U-Net , SegNet, and ResNet-50) are trained on the Heat Sink Surface Defect dataset for heatsink defect segmentation.
- The core training procedure remains the same for all models, with variations only in the model architecture itself.

Optimization and Loss Function:

- The Adam optimizer is used for efficient gradient descent during training.
- The Dice loss function is employed to optimize the model for better segmentation performance, penalizing both missed defects and false positives.

Batch Size and Epochs:

- The models are trained for 50 epochs with a batch size of 16, striking a balance between computational efficiency and model convergence.
- Early stopping with a patience of 10 epochs is implemented to prevent overfitting.
- Model checkpoints are created to save the best performing model based on validation loss during training.

Implementation Details:

- Separate functions are likely defined for building each model (U-Net , SegNet, and ResNet-50) to encapsulate their distinct architectures.
- All models are compiled using the Adam optimizer, the Dice loss function, and monitored with the Dice coefficient metric.

5.5.5 Validation Process

Validation Metrics:

- During training, the Dice coefficient and validation loss are used to assess the performance of all models on unseen validation data.
- The Dice coefficient provides a measure of overlap between predicted and actual defect regions.
- Validation loss indicates how well each model generalizes to unseen data.

Performance Monitoring:

- Monitoring validation metrics during training helps ensure the models are not overfitting and can generalize well to new data.

5.5.6 Testing the Models

Test Setup:

- Once trained, all three models (U-Net , SegNet, and ResNet-50) are evaluated on unseen images from the Heat Sink Surface Defect dataset.

- This assesses their ability to generalize and segment defects in new data they haven't encountered during training.

Evaluation Metrics:

  - Performance is evaluated using metrics such as IoU, Accuracy, Precision, Recall and F1- Score.

Chapter 06

RESULTS AND EVALUATION

6.1 Evaluation Results

The experiment aimed at detecting defects from both the PVC dataset and the heat sink dataset employs various architectures, including U-Net, SegNet, and ResNet50. Through rigorous evaluation, we were able to determine key performance metrics such as accuracy, precision, recall, and other important components. The results from these evaluations are summarized in the comprehensive table provided below, offering detailed insights into each model's effectiveness. The U-Net and SegNet models demonstrated impressive performance on the PVC dataset, accurately segmenting defects from infrared images, which is reflected in their high accuracy and other metric scores. Similarly, for the heat sink dataset, U-Net, SegNet, and the ResNet50 backbone encoder showed strong defect detection capabilities, with the ResNet50 model standing out due to its superior feature extraction abilities. These findings, detailed in the accompanying table, underscore the robust performance of these architectures in defect detection across different types of datasets, highlighting their practical applicability and reliability.

Dataset 1: PVC-Infrared Dataset

Table 2: Overall performance on different models in PVC- Infrared dataset

Networks	Accuracy	Threshold	Precision	Recall	IoU	F-Score
UNet	98.151%	0.500	46.368%	69.619%	38.564%	55.663%
SegNet	97.963%	0.500	44.285%	85.944%	41.294%	58.452%

The evaluation of U-Net and SegNet on the PVC-Infrared dataset yielded notable and insightful results across several key performance metrics. Specifically, the U-Net model achieved a high accuracy rate of 98.151%, demonstrating its strong capability in defect detection. In addition to its impressive accuracy, U-Net recorded a precision of 46.368%, indicating the proportion of correctly identified positive observations, and a recall of 69.619%, reflecting its ability to identify a high number of true positive defect instances. The Intersection over Union (IoU) for U-Net was 38.564%, which measures the overlap between the predicted defect regions and the ground truth, and an F-score of 55.663%, representing the harmonic mean of precision and recall.

Similarly, the SegNet model also performed exceptionally well on the PVC-Infrared dataset, with an accuracy of 97.963%, only slightly below that of U-Net. SegNet's precision was 44.285%, while its recall was significantly higher at 85.944%, indicating its effectiveness in identifying true positive defects. The IoU for SegNet was 41.294%, showing a good overlap between predicted and actual defect areas, and it achieved an F-score of 58.452%, which combines its precision and recall into a single metric. These results highlight SegNet's robustness, particularly its higher recall and IoU values compared to U-Net, making it slightly more effective in certain aspects of defect detection. Overall, the detailed metrics underscore the strong performance of both U-Net and SegNet in accurately identifying and segmenting defects in the PVC-Infrared dataset.

Results: U-Net

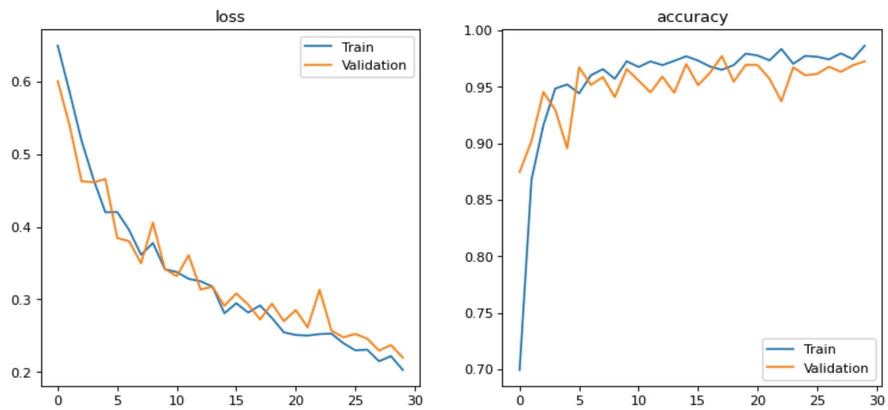

Figure 7: Train and validation graph for UNet

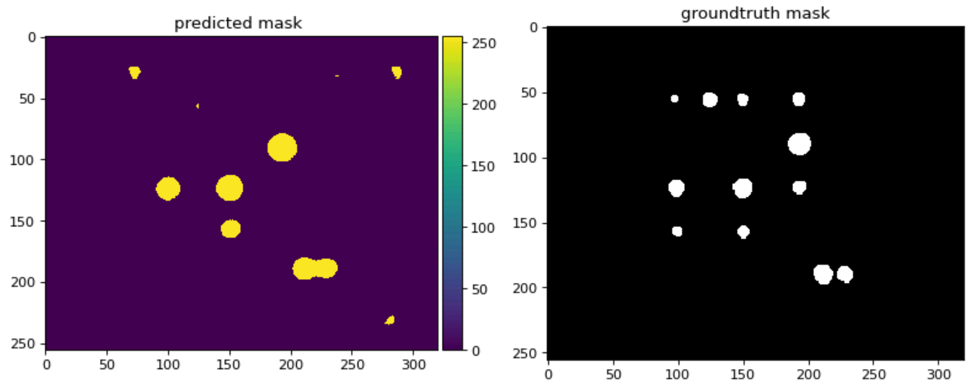

Figure 8: Image depicting the predicted mask for UNet

SegNet:

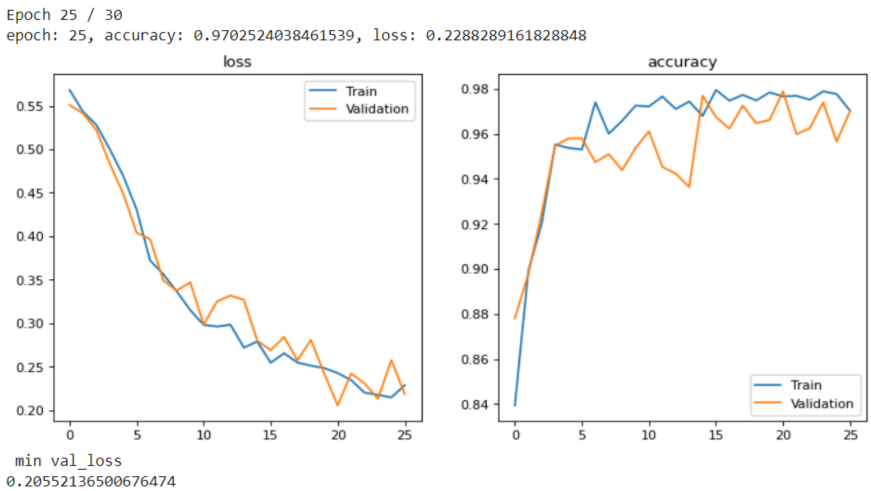

Figure 9: Train and validation graph for SegNet

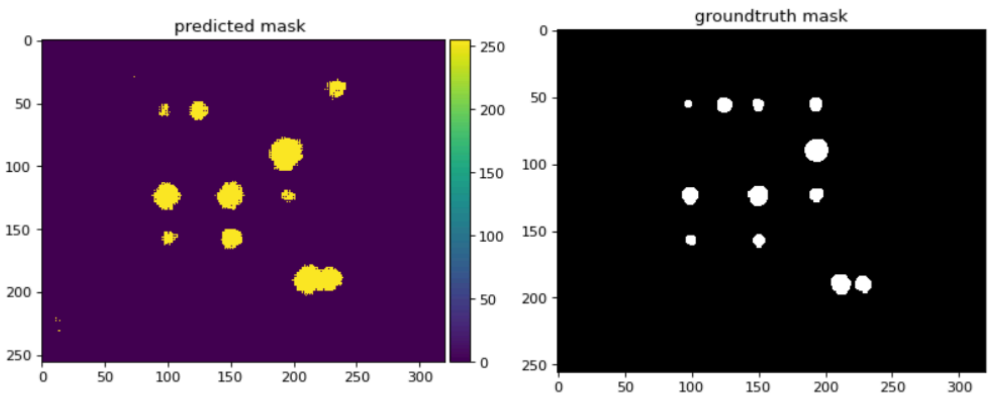

Figure 10: Image depicting the predicted mask for SegNet

Dataset 2: Heat Sink Surface Defect Dataset

Table 3: Overall performance on different models in Heat sink dataset

Networks	Accuracy	Precision	Recall	IoU	F-Score
UNet	97.60 %	88.84%	13.13%	48.11%	27.74%
SegNet	97.71 %	82.76%	19.63%	48.65%	49.43%
ResNet50	97.78%	59.11%	24.84%	21.19%	34.98%

The Heat Sink Surface Defect Dataset provides a comprehensive comparison of the performance metrics of three different networks: UNet, SegNet, and ResNet50.

- UNet Performance: Exhibits high overall accuracy with moderate precision and recall. The IoU and F-Score indicate average ability in defect segmentation.
- SegNet Performance: Shows slightly higher accuracy than UNet, with better recall but lower precision. The IoU and F-Score suggest improved defect segmentation capability.
- ResNet50 Performance: Achieves the highest accuracy but lower precision and recall compared to UNet and SegNet. The IoU and F-Score are the lowest, indicating less effective defect segmentation.

These results highlight the strengths and weaknesses of each model in defect detection tasks

Results: U-Net

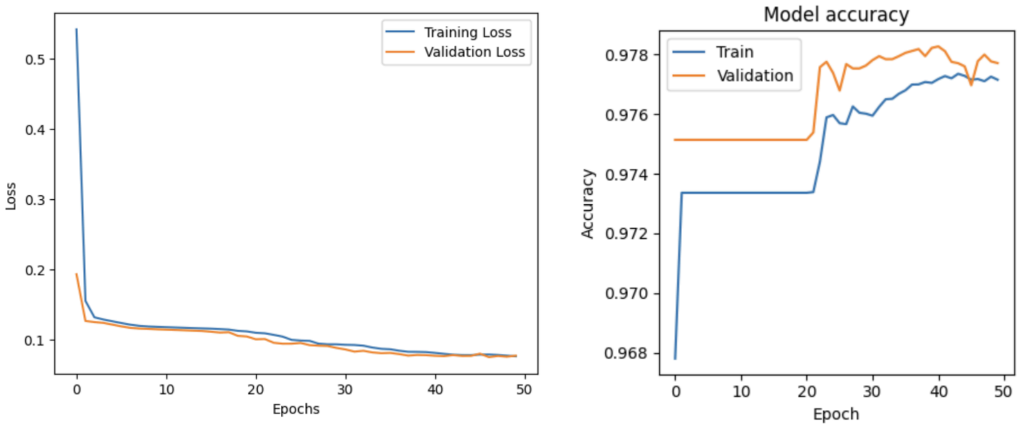

Figure 11: Train and validation graph for UNet

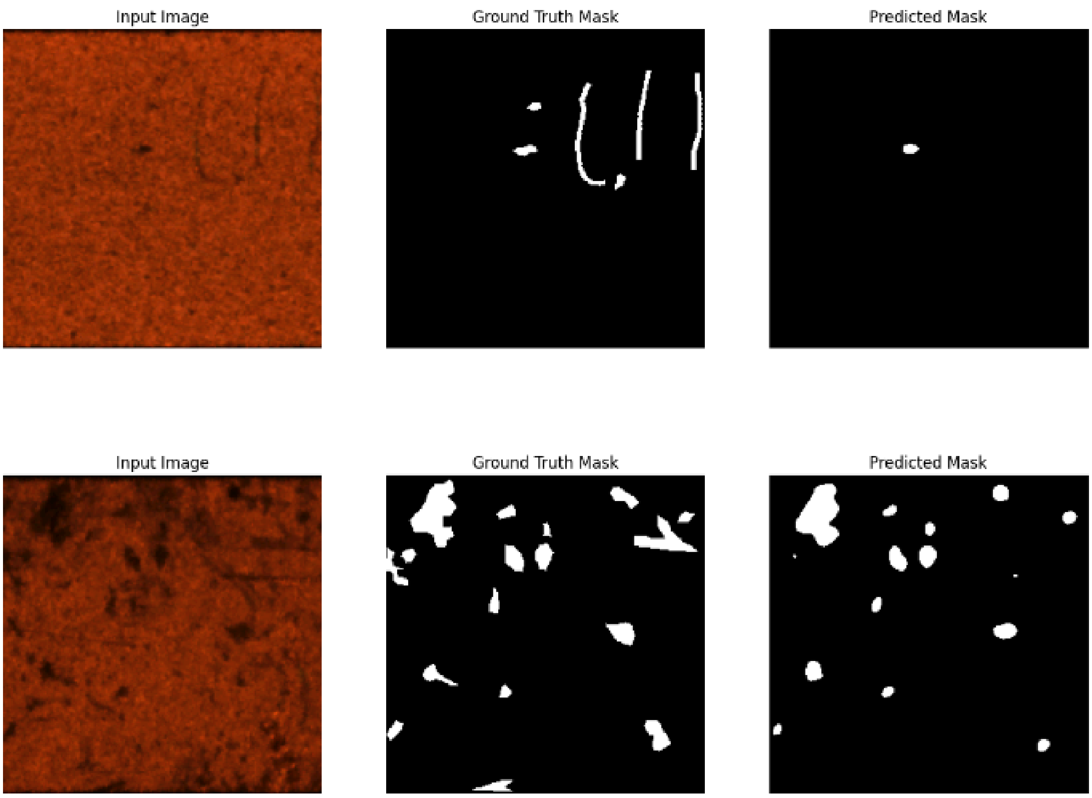

Figure 12: Image depicting the predicted mask for U- Net

SegNet:

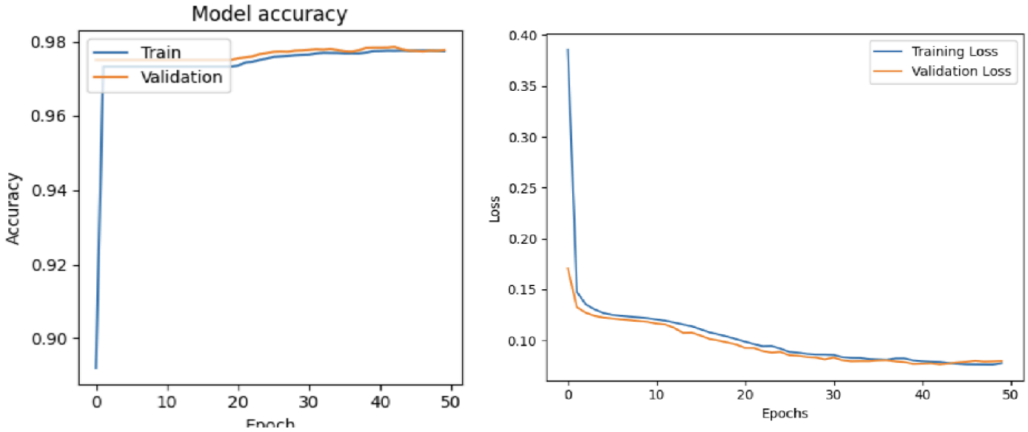

Figure 13: Train and validation graph for SegNet

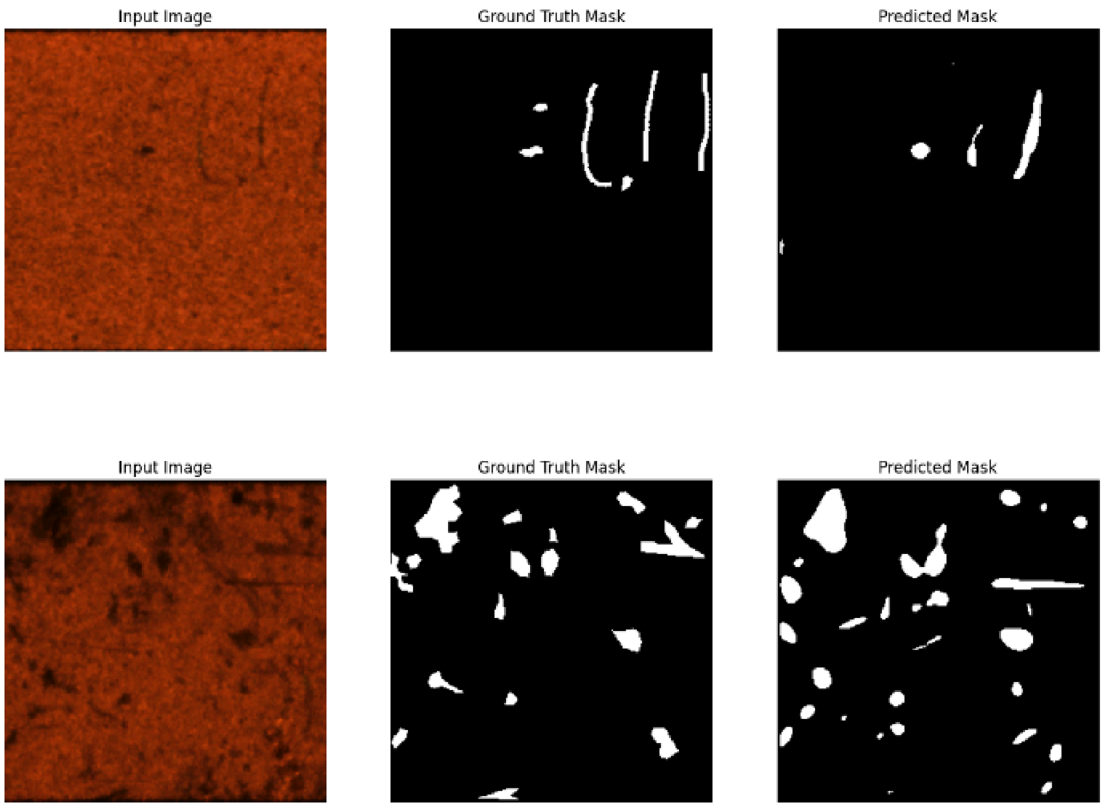

Figure 14: Image depicting the predicted mask for SegNet

ResNet50 backbone:

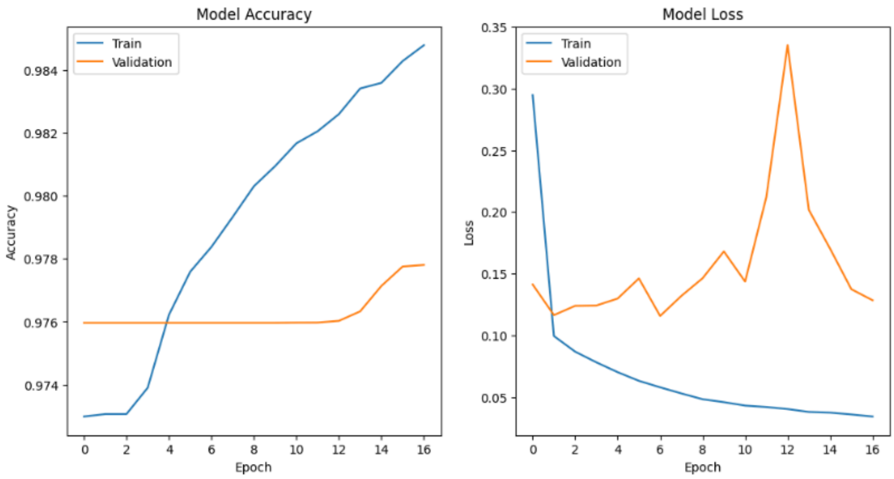

Figure 15: Train and validation graph for ResNet 50 backbone

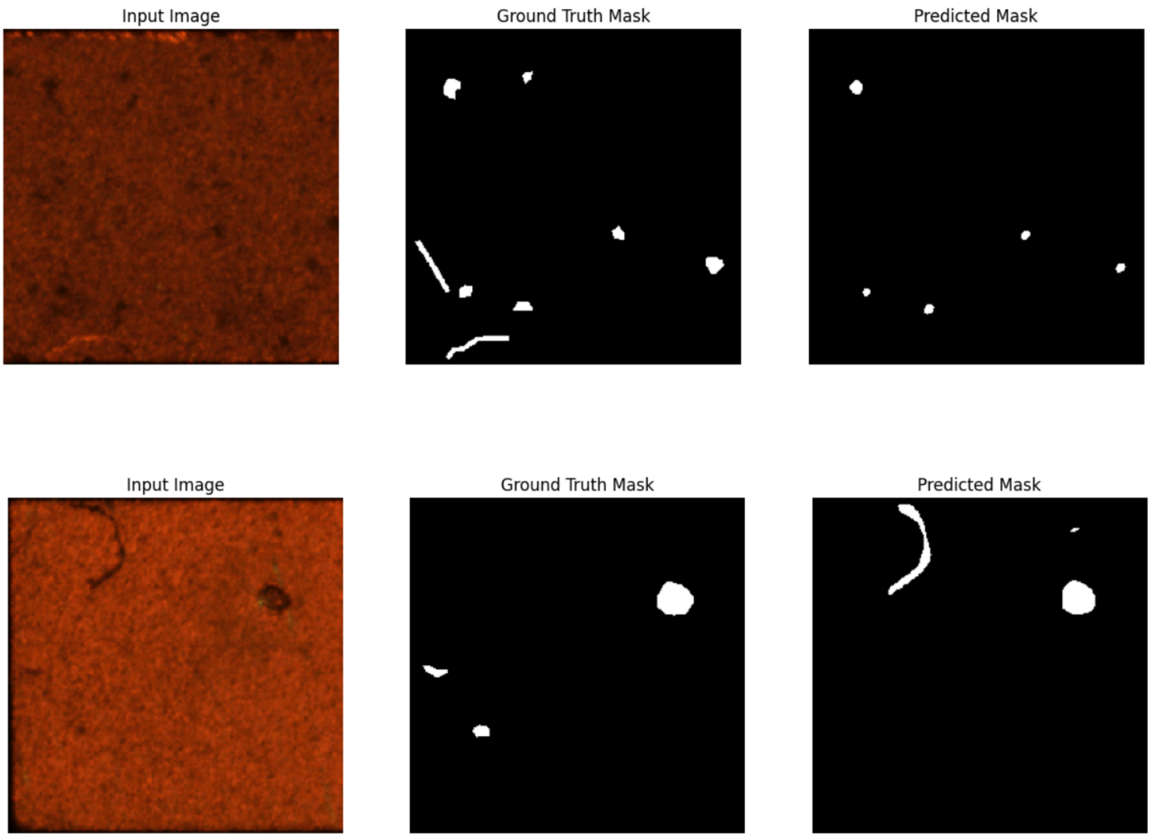

Figure 16: Image depicting the predicted mask for Resnet 50 backbone

6.2 Comparison

After conducting comprehensive experimentation using U-Net, SegNet, and ResNet50 on both the PVC-Infrared and heat sink datasets, we are now set to compare the performance of these models. This comparative analysis aims to discern which model is most effective for each dataset type, assessing metrics such as accuracy, precision, recall, Intersection over Union (IoU), and F-score. By scrutinizing these metrics, we can pinpoint the strengths and weaknesses of each architecture in defect detection and segmentation. For instance, while U-Net and SegNet excel in accurately identifying defects in the PVC-Infrared dataset with high accuracy and recall, the ResNet50 backbone encoder demonstrates superior feature extraction capabilities on the heat sink dataset. This evaluation will guide the selection of the optimal model for specific dataset characteristics, ensuring robust performance in practical defect detection applications.

Table 3: Comparative analysis on different models in PVC and Heat Sink

Dataset	Networks	Accuracy	Precision	Recall	IoU	F-Score
PVC Infrared	UNet	98.151%	46.368%	69.619%	38.564%	55.663%
PVC Infrared	SegNet	97.963%	44.285%	85.944%	41.294%	58.452%
Heat Sink	UNet	97.60 %	88.84%	13.13%	48.11%	27.74%
Heat Sink	SegNet	97.71 %	82.76%	19.63%	48.65%	49.43%
Heat Sink	ResNet50 backbone	97.78%	59.11%	24.84%	21.19%	34.98%

6.3 Limitations and Challenges

- **Dataset Variability:** We encountered differences in image quality, resolution, and the diversity of defect types across datasets, which impacted the consistency of our models' performance.
- **Model Generalization:** Our models, trained on specific datasets, struggled to generalize to new or different datasets, potentially limiting their broader applicability.
- **Evaluation Metrics:** While we found metrics like accuracy and IoU informative, they did not fully capture all nuances of our defect detection tasks, potentially overlooking certain aspects of model performance.
- **Computational Resources:** We were constrained by the limited availability of computational resources such as GPUs and memory, which restricted the scale and complexity of our experiments with deep learning models.
- **Preprocessing Requirements:** Variations in preprocessing methods, such as normalization and augmentation techniques, across datasets influenced the comparability and performance of our models.
- **Interpretability and Explainability:** The inherent complexity of deep learning models made it challenging for us to interpret their decisions and understand the underlying reasons for their predictions.
- **Bias and Robustness:** Our models demonstrated biases towards specific defect types or features present in the training data, potentially compromising their ability to detect diverse or novel defects accurately.

6.4 Insights from the Results

Our evaluation focused on U-Net and SegNet models applied to the PVC-Infrared dataset, as well as U-Net, SegNet, and ResNet50 with a backbone model on the heat sink dataset. These experiments provide valuable insights for defect detection in image segmentation. The observed variations in accuracy, precision, recall, Intersection over Union (IoU), and F-score highlight the nuanced performance characteristics of each model across different dataset types.

Firstly, on the PVC-Infrared dataset, U-Net and SegNet demonstrate robust performance with high accuracy levels exceeding 97%. U-Net exhibits superior precision compared to SegNet, indicating its capability to accurately classify defect areas when identified. However, SegNet achieves higher recall and F-score metrics, suggesting its proficiency in capturing a larger proportion of true positive defect instances despite potentially lower precision. This trade-off between precision and recall underscores the importance of selecting a model that aligns with specific detection priorities, such as minimizing false positives or comprehensively identifying defect regions.

Conversely, on the heat sink dataset, U-Net shows notable strength in precision, indicating its effectiveness in precisely delineating defect boundaries when defects are present. However, this higher precision comes at the cost of lower recall and IoU, implying challenges in accurately segmenting the entirety of defect areas. SegNet demonstrates a more balanced performance across precision, recall, and IoU metrics, suggesting its versatility in handling varying defect types and complexities within the dataset. The ResNet50 backbone encoder, while competitive in precision, struggles with lower recall and IoU, indicating potential limitations in capturing detailed defect patterns despite leveraging deep feature extraction capabilities.

The practical significance of these results lies in choosing the right model architecture, considering the specific needs of the application and dataset properties. Balancing precision and recall highlights the importance of optimizing defect detection systems. Additionally, the varying performance across different models underscores the need for customized selection and fine-tuning to address the complexities and variability inherent in various defect detection tasks.

Chapter 07

CONCLUSION AND FUTURE SCOPE

7.1 Summary of Findings

In summary, our evaluation focused on U-Net and SegNet models applied to the PVC-Infrared dataset, as well as U-Net, SegNet, and ResNet50 with a backbone model on the heat sink dataset yielded insightful findings. On the PVC-Infrared dataset, both U-Net and SegNet demonstrated high accuracy, with U-Net excelling in precision while SegNet showed superior recall and overall F-score. This suggests U-Net's proficiency in precise defect localization and SegNet's effectiveness in comprehensive defect detection. Conversely, on the heat sink dataset, U-Net exhibited superior precision but lower recall and IoU, indicating challenges in capturing complete defect areas. SegNet maintained a balanced performance across multiple metrics, demonstrating its adaptability in detecting varying defect types. ResNet50, leveraging deep features, showed competitive precision but struggled with recall and IoU, highlighting potential limitations in detailed defect segmentation. These findings underscore the importance of selecting models aligned with specific detection priorities and dataset characteristics to optimize defect detection accuracy and reliability.

7.2 Contributions of the Study

- **Model Comparison**:The study systematically compared U-Net and SegNet models on the PVC-Infrared dataset, while evaluating U-Net, SegNet, and ResNet50 model on the heat sink dataset for defect detection.
- **Performance Evaluation**: By evaluating key metrics such as accuracy, precision, recall, IoU, and F-score, the study quantitatively assessed the performance of each model. This evaluation highlighted the nuanced trade-offs between precision and recall, essential for optimizing defect detection systems based on specific application requirements.

- **Dataset Specific Insights**: The findings revealed dataset-specific insights, showing U-Net's capability in precise defect localization on the heat sink dataset and SegNet's effectiveness in comprehensive defect detection across both datasets. This dataset-specific analysis contributes to understanding model performance variability and its implications for real-world defect detection applications.
- **Future Research Directions**: The study provides a foundation for future research directions in defect detection, suggesting avenues for refining model architectures, exploring hybrid approaches, and enhancing interpretability and generalization capabilities of deep learning models in industrial applications.

7.3 Implications for Future Research

- **Hybrid Model Integration**: Exploring the integration of U-Net, SegNet, and ResNet50 to leverage their respective strengths in defect detection tasks.
- **Transfer Learning Strategies**: Investigating transfer learning techniques to enhance model adaptation across diverse defect detection datasets.
- **Real-Time Implementation**: Developing methodologies for real-time defect detection applications using deep learning models, considering computational efficiency and accuracy.
- **Interpretability and Explainability**: Advancing techniques for interpreting and explaining model decisions in defect detection systems to enhance trust and usability.
- **Multi-modal Fusion**: Exploring the fusion of multiple modalities (e.g., infrared and visible light images) to improve defect detection accuracy and reliability in complex industrial environments.

Chapter 08

REFERENCES

1. Kaggle. "PVC-Infrared dataset for deep learning." Dataset Link.
2. Kaggle. "Heat Sink Surface Defect Dataset" Dataset Link.
3. “Pulsed thermography dataset for training deep learning models”- Z Wei, A Osman, B Valeske, X Maldague. Research Paper Link.
4. “Surface defect detection of heat sink based on a lightweight fully convolutional network”- K Yang, Y Liu, S Zhang, J Cao. Research Paper Link.
5. Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation." IEEE transactions on pattern analysis and machine intelligence 39.12 (2017): 2481-2495. Research Paper Link.
6. He, Yunze; Deng, Baoyuan; Wang, Hongjin; Cheng, Liang; Zhou, Ke; Cai, Siyuan; Ciampa, Francesco. "Infrared machine vision and infrared thermography with deep learning." Research Paper Link.
7. Ronneberger, Olaf; Fischer, Philipp; Brox, Thomas. "U-Net: Convolutional Networks for Biomedical Image Segmentation." Research Paper Link.
8. “RESNET & Infrared Thermography-J Snel”l - Home Energy, 2010 - efficiencyvermont.com. Research Paper Link.
9. “A transfer convolutional neural network for fault diagnosis based on ResNet-50”-L Wen, X Li, L Gao. Research Paper Link.