Realistic Image Synthesis and Enhancement Using Deep Learning Architecture for Low-Resource Image Recognition Tasks

Shivam Samarendu Chaudhary
Master Of Engineering, Department of Computer Engineering,
Shankersinh Vaghela Bapu Institute of Technology

Abstract
Image recognition systems operating in low-resource environments often face challenges arising from poor lighting, low resolution, image degradation, and limitations of image-capturing hardware. Such conditions can adversely affect the reliability of automated recognition and detection systems. In automobile inspection, these challenges become particularly important because vehicle damage such as dents, cracks, and other defects may be captured through images of varying quality and under challenging imaging conditions.
This Dissertation Phase – II presents an experimental study on a hybrid deep learning Model that integrates a Generative Adversarial Network (GAN) for realistic image synthesis and image enhancement with the You Only Look Once (YOLO) architecture for automobile damage detection. The study extends the research direction established during Dissertation Phase – I and evaluates the proposed approach using the Automobile Damage Detection dataset obtained from Roboflow Universe. The dataset comprises 6,957 annotated images belonging to fourteen automobile damage classes. The proposed Model involves image preprocessing and standardisation, GAN-based image enhancement and realistic synthetic image generation, followed by YOLO-based object detection and classification.
The developed hybrid Model was experimentally evaluated using accuracy, precision, recall, and F1-score. The results demonstrated that the hybrid YOLO–GAN model achieved an accuracy of 96%, precision of 93%, recall of 91%, and F1-score of 92%. The results demonstrate the effectiveness of integrating GAN-based image enhancement and synthesis with YOLO-based object detection for improving automobile damage recognition under challenging and low-resource imaging conditions.
The study establishes a hybrid deep learning approach for automobile damage detection and demonstrates the potential of combining generative and detection-based architectures within a unified recognition Model. The proposed approach can support applications including automobile inspection, insurance claim assessment, vehicle service inspection, and roadside damage assessment. Future work may focus on computational optimisation, lightweight architectures, and deployment in resource-constrained environments.

Keywords: Image Synthesis, Image Enhancement, Generative Adversarial Network, YOLO, Hybrid Deep Learning, Automobile Damage Detection, Low-Resource Image Recognition, Object Detection, Data Augmentation, Deep Learning.

CHAPTER 1 - INTRODUCTION

1.1 Research Background

Image recognition has become a core capability of contemporary artificial intelligence systems across diverse industries such as healthcare diagnostics, autonomous navigation, security surveillance, agriculture, and digital media. The reliability of an image recognition system, however, is highly influenced by the input image's quality. In practical use cases, images are acquired under low-resource conditions involving low resolution, inadequate illumination, noise, occlusion, and limitations of image-capturing hardware. Such conditions can adversely affect the extraction of meaningful visual features and consequently reduce the reliability of recognition and detection systems. The challenges associated with low-quality images and limited training data formed the foundation of the research proposed during Dissertation Phase – I.

Dissertation Phase – I proposed a hybrid model integrating a GAN - Generative Adversarial Network for image enhancement and realistic synthetic image generation along with the YOLO - You Only Look Once architecture for object detection and classification in real-time. The proposed research was designed around an integrated workflow consisting of data input, preprocessing, GAN-based enhancement and synthesis, YOLO-based detection, and performance evaluation. The underlying objective was to address the limitations of approaches that treat image enhancement, synthetic data generation, and object detection as separate processes, particularly in low-resource environments.

Dissertation Phase – II is researched deeply with experiments and evaluation of the hybrid model for automobile damage detection. The application focus is refined from general image recognition toward the identification and localisation of specific automobile damage categories, including dents and damage to vehicle components. The experimental study uses the Automobile Damage Detection dataset published on Roboflow Universe, comprising 6,957 annotated images across fourteen automobile damage classes. Phase - II provides a deeper understanding and evaluation of the proposed model and examines its effectiveness for automobile damage recognition under challenging image-recognition conditions.

1.2 Problem Context

Automobile damage assessment involves the identification and recognition of visible damage to vehicle components through images. In real-world image recognition environments, the quality of the input image can vary considerably due to factors such as low resolution, inadequate lighting, noise along with limitations of image-capturing hardware that can make it difficult for automated recognition systems to extract and identify the relevant visual features accurately.

The problem becomes more significant when image enhancement, synthetic data generation, and object detection are addressed as separate processes. Image enhancement may improve visual quality but does not necessarily ensure improved detection performance, while object detection systems are strongly dependent on the quality of their input images. Similarly, conventional data augmentation techniques may increase dataset size but have limited ability to generate realistic and diverse new samples. These limitations create a need for an integrated approach that can improve image quality, generate additional realistic training data, and perform reliable object detection within a unified framework.

1.3 Problem Statement

1. Lack of an Integrated System:

Existing approaches generally address image enhancement, image synthesis, and object detection as separate tasks. There is a need for an integrated framework capable of combining these functions within a unified deep learning pipeline.

2. High Computational Cost:

Advanced image enhancement, generative, and detection models can involve considerable computational requirements. High computational cost creates challenges for their practical use in resource-constrained environments and highlights the need for an approach that balances performance with computational efficiency.

3. Performance Degradation with Low-Quality Images:

The effectiveness of image recognition and object detection models can decrease when the input images suffer from low resolution, noise, inadequate illumination, or other forms of image degradation. Such conditions can limit the availability of useful visual information required for reliable detection.

4. Limited Utilisation of Synthetic Data:

Although generative models such as GANs can be used to generate realistic synthetic images, the potential of generated data is not always effectively incorporated into the object-detection process. This limits the utilisation of synthetic data for improving the diversity and availability of training samples.

5. Absence of an End-to-End Solution:

There is a need for a comprehensive end-to-end approach that can collectively enhance image quality, generate additional realistic data, and perform object detection within a unified framework. The absence of such an integrated solution forms the central research problem addressed by the proposed GAN–YOLO Model.

1.4 Motivation

The motivation for this research arises from the need to improve the reliability and practical applicability of deep learning-based image recognition under low-resource conditions. Existing approaches may perform effectively when adequate image quality and computational resources are available; however, challenges associated with low-resolution images, noise, inadequate illumination, and limited computational resources can affect their effectiveness.

A further motivation is the limited integration of synthetic data generation with image recognition and object detection processes. Generative models provide an opportunity to enhance image quality and generate realistic synthetic samples, while object detection models provide the capability to identify and classify relevant objects. Combining these capabilities can therefore provide a more comprehensive approach than treating them as independent processes.

The research is consequently motivated by the need to develop a robust, resource-conscious, and integrated deep learning approach that can enhance image quality, utilise realistic synthetic data, and improve object detection performance within a unified framework. The proposed GAN–YOLO approach is intended to address these requirements and provide a foundation for practical image-recognition applications in low-resource environments.

1.5 Need for the Research

The need for this research arises from the following key requirements:

- To improve the visual quality of degraded automobile images while preserving the important visual features required for reliable damage detection.
- To generate realistic synthetic images to increase the diversity of available training data and support recognition of different automobile damage categories.

- A method is needed that can balance recognition performance with computational requirements so that the resulting approach remains suitable for resource-constrained environments.
- A system that can integrate image enhancement, synthetic image generation, and object detection within a unified framework rather than treating these processes independently.
- To measure the efficacy of the proposed hybrid model for automobile damage detection under low-resource image-recognition conditions.

1.6 Aim of the Research

The primary aim of this study is to achieve the highest accuracy by designing, implementing, and experimentally evaluating a hybrid deep learning Model that integrates pre-processing of data followed by GAN-based image enhancement and synthetic image generation with YOLO-based object detection to improve the efficacy of Automobile Damage Recognition under low-resource imaging conditions, using the Automobile Damage Detection dataset as the basis for evaluation.

1.7 Research Objectives

The primary objectives of the research are to:

1. Design a GAN-based module for enhancing low-quality automobile images and generating realistic synthetic image samples.
2. Adapt and fine-tune a YOLO-based detector for multi-class automobile damage detection.
3. Integrate GAN-based image enhancement and synthetic image generation with YOLO-based object detection within a unified hybrid Model.
4. Experimentally evaluate the efficacy and robustness of the proposed hybrid Model for automobile damage detection using appropriate performance measures.
5. Examine the computational requirements and practical applicability of the proposed Model in low-resource image-recognition environments.

1.8 Tools used for Proposed System

The following tools were used for the implementation of the proposed system:

- **Programming Language**
  - Python: The primary programming language used due to its simplicity and massive libraries.
- **Deep Learning Model**
  - PyTorch: Used for building and handling deep learning models, GAN and YOLO architectures.
- **GPU Acceleration**
  - CUDA: To speed up the training of deep learning models that significantly reduced the computation time.
- **Image Processing**
  - OpenCV: For Image preprocessing operations like resizing, filtering and augmentation
- **Datasets**
  - Roboflow: Automobile damage detection dataset used for training and testing

1.9 Major Contribution of the Research

The key contributions of the research include:

1. **Development of proposed Hybrid Model**integrating GAN-based image enhancement and synthetic image generation with YOLO-based automobile damage detection.
2. **Experimental investigation of the proposed hybrid approach** for multi-class automobile damage recognition using a publicly available automobile damage dataset.
3. **Evaluation of the hybrid model under standard measures for performance**, including accuracy, precision, recall, and F1-score.
4. **Investigating the hybrid model's application to low-resource image-recognition conditions**, with consideration of image quality, synthetic data utilisation, and computational requirements.
5. **Establishment of an integrated Model** combining image enhancement, synthetic image generation, and object detection as proposed in the research direction established during Dissertation Phase–I.

CHAPTER 2 - COMPREHENSIVE LITERATURE REVIEW

2.1 Introduction

This chapter is a comprehensive review of the research on the hybrid GAN–YOLO model. The review covers major developments in deep learning-based image recognition, image quality and low-resource recognition, conventional and deep learning-based image enhancement, GAN-based image synthesis, advanced generative models, synthetic data generation, object detection architectures, YOLO-based detection, and hybrid approaches combining generative and detection models.

2.2 Deep Learning for Image Recognition

Deep learning has significantly advanced image recognition by enabling models to learn hierarchical and discriminative visual features directly from image data. In contrast to traditional approaches that rely extensively on manually designed feature descriptors, deep learning architectures can learn increasingly complex representations through multiple layers of computation. This capability has contributed to the adoption of deep learning across a wide range of image-processing and visual-recognition applications.

Hidayat et al. (2025), through a bibliometric analysis of artificial intelligence in digital image processing, highlight the increasing research activity and application of AI-based approaches in image-processing tasks. The study reflects the broader development of intelligent image-analysis systems and the growing role of deep learning in automated visual processing.

Object recognition has also benefited from the availability of synthetic and diverse training data. Baimukashev et al. (2019) investigate "Deep learning-based object recognition using physically realistic synthetic depth scenes", demonstrating the relevance of synthetic data in supporting recognition tasks where suitable real-world data may be limited. This provides an important foundation for considering synthetic image generation as a component of a broader recognition pipeline.

More recently, transformer-based architectures have emerged as an important development in computer vision. Hassija et al. (2025) review transformer-based approaches for vision and discuss their ability to model relationships between image features over broader spatial contexts. These developments demonstrate that image recognition has progressed beyond conventional convolutional architectures and now includes multiple architectural approaches with different computational and representational characteristics.

Despite these advances, the effectiveness of an image recognition system remains dependent not only on the architecture used but also on the characteristics and quality of the available input data. Consequently, improving image quality and increasing the diversity of training data remain important considerations for reliable recognition, particularly when images are acquired under challenging conditions.

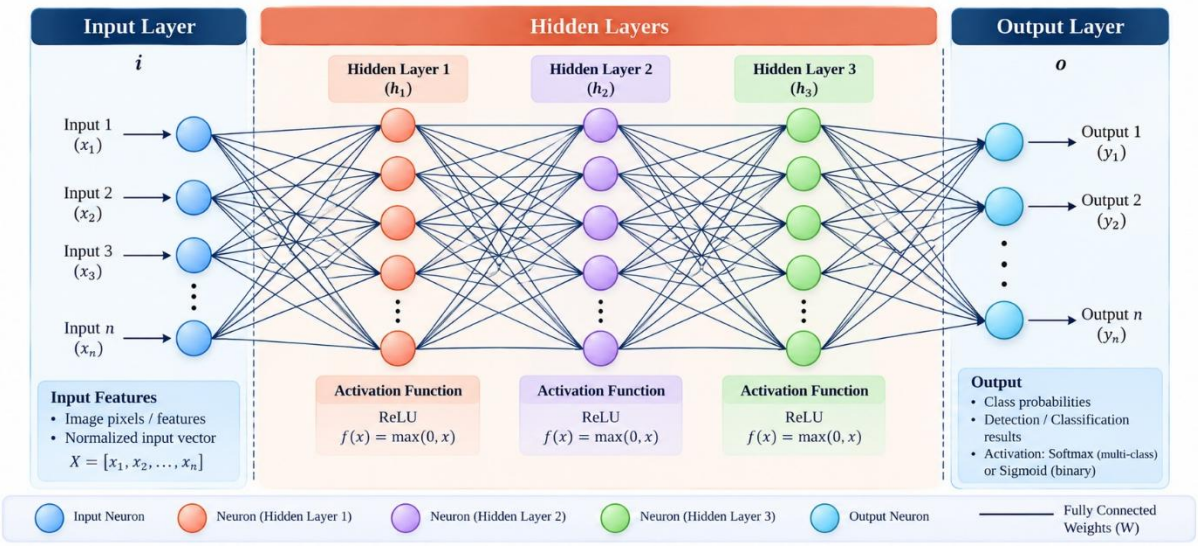


Figure 2.1: Deep Learning for Image Recognition

2.3 Image Quality and Low-Resource Recognition

The quality of the input image is an important factor in image recognition and detection performance. Real-world images may contain variations in illumination, resolution, noise, contrast, viewing conditions, and other acquisition-related characteristics. Such variations can affect the visual information available to a recognition model and consequently influence its ability to identify relevant features.

Low-light imaging represents one important example of this challenge. The literature reviewed during Phase–I indicates that deep learning-based enhancement methods can improve the visibility and contrast of low-light images while attempting to preserve features relevant to subsequent recognition tasks. The reviewed work also highlights the importance of feature preservation, since excessive enhancement may introduce artifacts or amplify existing noise. A recurring theme across the literature is that recognition performance degrades sharply once input images deviate from the clean, well-lit conditions typical of curated benchmarks. Asha and Hemalatha review deep learning approaches for enhancing low-light photography; Chen et al. (2023) survey image enhancement and correction methods in medical imaging, noting that illumination variation alone can substantially depress downstream recognition accuracy. Farahzadi et al. (2025), in the specific context of built-environment condition assessment, catalogue the damage characteristics and detection requirements that make automated visual inspection difficult when source imagery is inconsistent in resolution, angle, and lighting — conditions that mirror those encountered in automobile damage photography.

The problem becomes particularly relevant in low-resource recognition environments, where limitations may arise from both the quality of the available images and the computational resources available for processing them. High-quality input data may not always be available, while advanced enhancement and recognition architectures can themselves require substantial computational resources.

These considerations establish two important requirements for the present research: first, the ability to improve or enhance image information before detection; and second, the ability to utilise additional realistic image data to support the recognition process. These requirements provide the basis for examining image enhancement, GAN-based synthesis, and object detection as interconnected components of the proposed hybrid Model.

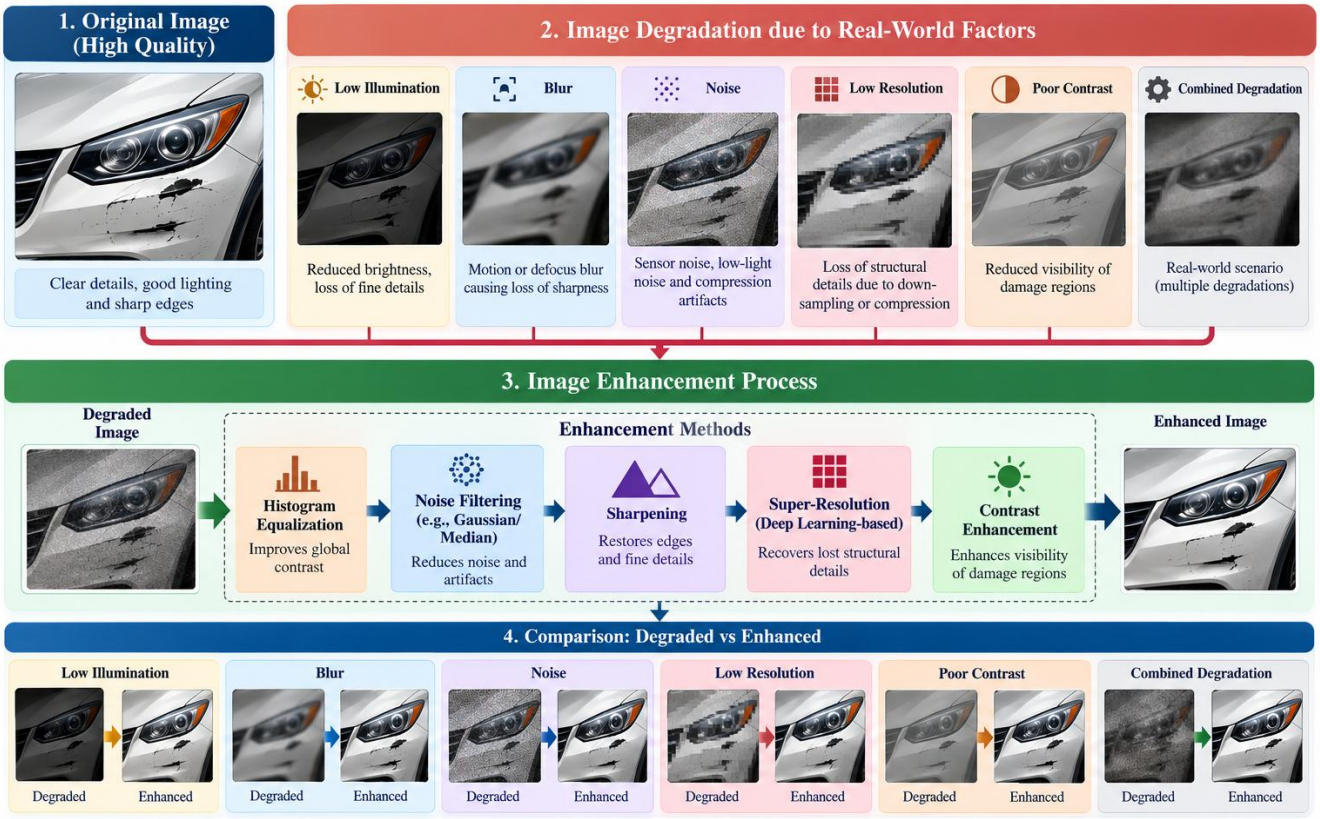


Figure 2.2: Image Quality Degraded vs Enhanced

2.4 Conventional Image Enhancement Techniques

The previous techniques used for image enhancement have traditionally been used to improve the quality of digital images before analysing them for further use. Common approaches include interpolation, histogram equalisation, contrast adjustment, and classical filtering methods. These techniques can improve brightness, contrast, sharpness, or apparent image quality without requiring the training of a dedicated deep learning model.

Such methods are generally computationally simple and can therefore be useful when processing resources are limited. However, their operations are primarily based on predefined mathematical transformations or local image characteristics rather than features learned from data. Consequently, their ability to restore information lost through degradation may be limited, particularly when images contain substantial noise, low resolution, poor illumination, or other acquisition-related distortions.

Another important consideration is the preservation of information relevant to subsequent recognition. Enhancement operations that excessively smooth an image or alter its contrast may affect fine visual structures that are useful for detection. In automobile damage recognition, such structures may include the boundaries of dents, surface irregularities, cracks, or other local damage characteristics.

The limitations of conventional enhancement techniques therefore provide a motivation for investigating learning-based approaches. Instead of applying fixed transformations, deep learning-based enhancement methods can learn relationships between degraded and higher-quality image representations from training data. This transition from manually designed enhancement operations to learned image restoration forms an important development in the literature.

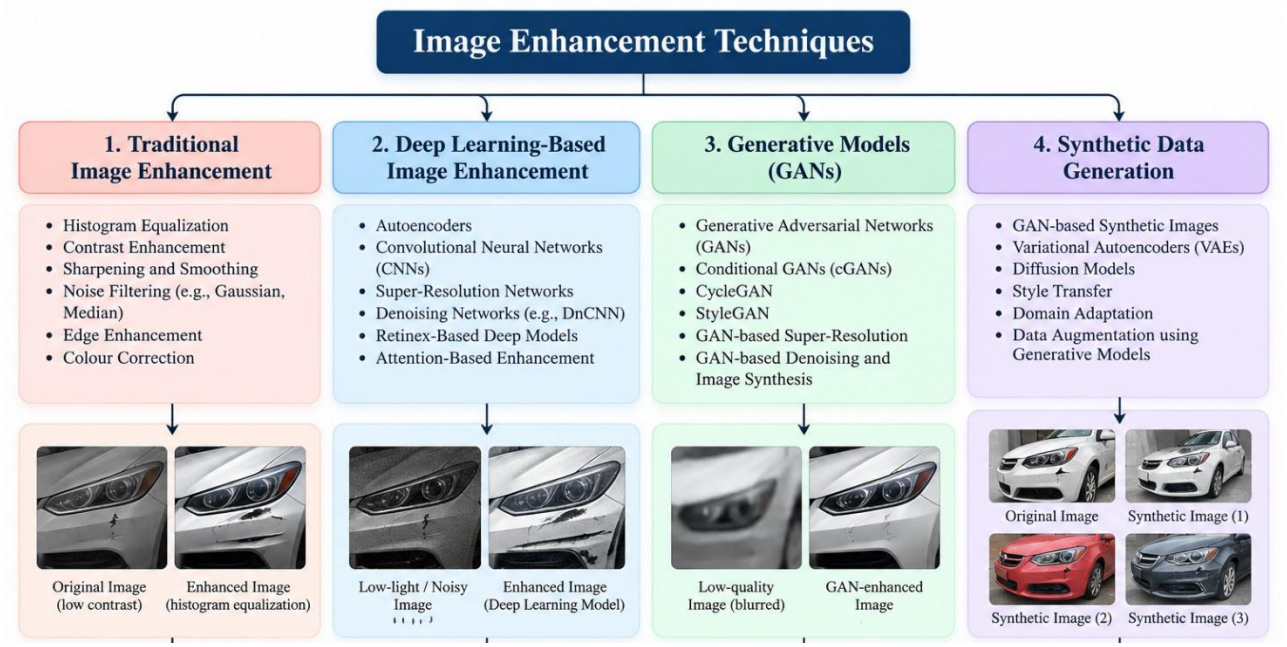


Figure 2.3: Image Enhancement Techniques

2.5 GAN-Based Image Synthesis

Goodfellow et al. (2014), introduced Generative Adversarial Networks (GANs) that was an important development in generative deep learning. It stated, 'GAN consists of two neural networks: a generator, which produces synthetic samples, and a discriminator, which attempts to distinguish generated samples from real samples.' An adversarial process was used to train these two networks, enabling the GAN generator to progressively improve the quality of its generated outputs over time.

The ability of GANs to learn the distribution of image data has made them applicable to several image-processing tasks, including image synthesis, image enhancement, super-resolution, style transformation, and data augmentation. Toshpulatov et al. (2021), in their survey of GANs and their applications to three-dimensional face generation, discuss the ability of GAN-based approaches to generate realistic and diverse visual samples. The broader literature reviewed during Phase–I similarly identifies high-fidelity synthetic image generation as an important application of GANs in situations where labelled or suitable training data may be limited.

GANs can be particularly relevant to image-recognition systems because synthetic samples can provide additional variation in the training data. When appropriately generated and incorporated into a training pipeline, such samples may help expose a recognition model to a broader range of visual characteristics. This makes GAN-based synthesis relevant to the present study, where the objective is to investigate the combination of image enhancement and synthetic image generation with object detection.

GANs are considered not only as independent image-generation models but also as a component of an integrated recognition pipeline. Their role is to support image enhancement and realistic synthetic image generation, after which the resulting visual information can be utilised by the YOLO-based detection component.

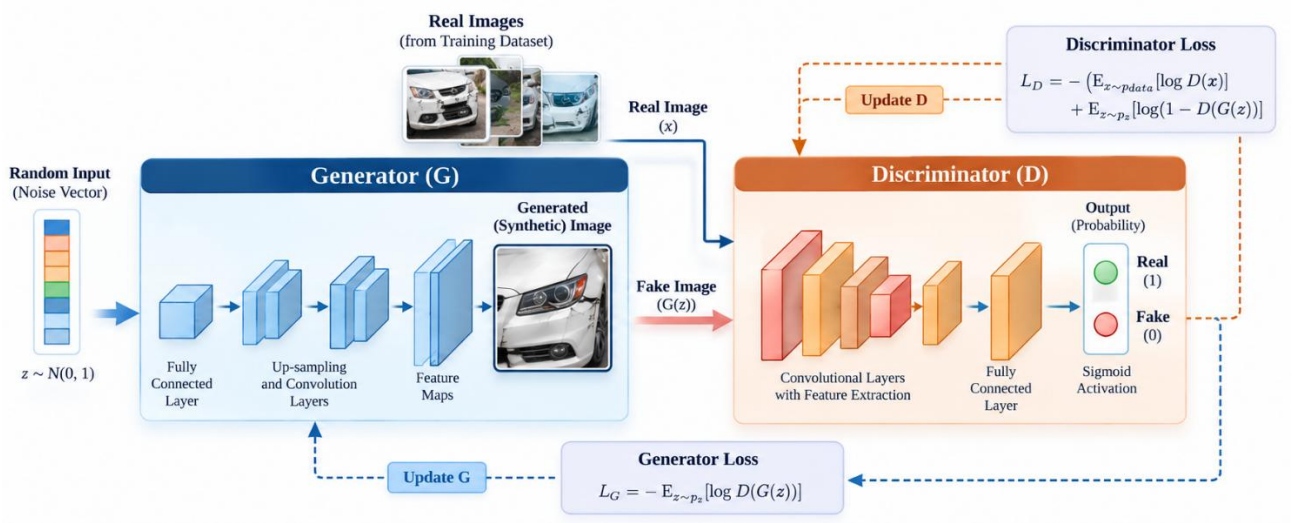


Figure 2.4: Generative Adversarial Network Model

2.6 Advanced Generative Models

The development of generative modelling has extended beyond conventional GAN architectures to include diffusion-based and other advanced generative approaches. These models provide alternative mechanisms for image synthesis and restoration and have demonstrated strong capabilities in generating visually realistic outputs.

Saharia et al. (2022) introduced SR3, a diffusion-based super-resolution approach that progressively refines an image through an iterative generation process. The reviewed literature identifies diffusion-based approaches as capable of producing highly detailed and photorealistic outputs. However, the iterative nature of diffusion-based generation can result in higher computational requirements and slower inference compared with approaches designed for faster generation.

Real-ESRGAN, proposed by Wang et al. (2021), extends the Enhanced Super-Resolution Generative Adversarial Network approach toward real-world image degradation. The method is designed to handle complex and unknown degradations rather than relying exclusively on idealised degradation models. The literature reviewed in Phase–I identifies Real-ESRGAN as an effective approach for practical image enhancement while also recognising its computational requirements.

GFP-GAN, proposed by Wang et al. (2021), uses generative priors for restoration of degraded facial images. Although its application is domain-specific, it demonstrates how generative priors can be incorporated into restoration systems to recover visually meaningful information.

These developments demonstrate that modern generative models can provide high-quality image restoration and synthesis capabilities. At the same time, the increasing complexity of advanced generative architectures creates an important trade-off between output quality and computational efficiency. This trade-off is particularly relevant to the present research because the proposed Model is intended to investigate image enhancement and synthesis in the context of low-resource image recognition.

2.7 Synthetic Data and Data Augmentation

The diversity of training data is an important consideration while developing deep-learning-based recognition systems. Traditional techniques of data augmentation like image rotation, cropping, scaling, flipping and colour modification increase the variation within a training dataset without the requirement of a completely new image dataset.

Beyond conventional GANs, diffusion-based generative models have emerged as an alternative for photorealistic image synthesis. Saharia et al. (2022) describe SR3, an iterative diffusion-based super-resolution model that progressively refines degraded inputs. Wang et al. (2021b) present Real-ESRGAN for real-world blind super-resolution, while GFP-GAN (Wang et al., 2021a) focuses on facial restoration using generative priors. Dunlap et al. (2023) introduce ALIA, a diffusion-based augmentation approach designed to diversify vision datasets. These methods broaden the options for generating or restoring images, although their computational requirements vary and may be relevant in low-resource settings.

The use of synthetic data nevertheless requires careful consideration. Generated samples must retain meaningful visual characteristics and should be sufficiently representative of the target domain to provide useful information to a recognition model. Poor-quality or unrealistic synthetic samples may provide limited benefit and can potentially introduce unwanted variation into the training process. For the present research, synthetic data generation is therefore considered as a component of the hybrid Model rather than as an independent objective. GAN-based synthesis is investigated alongside image enhancement and YOLO-based detection, with the intention of providing additional realistic visual information within the overall recognition pipeline.

2.8 Object Detection Architectures

Object detection involves both identifying the presence of objects within an image and determining their spatial locations, typically through bounding-box predictions. Modern object detection architectures have developed along several major directions, including two-stage and single-stage detection approaches.

Two-stage detectors generally separate the detection process into region proposal and subsequent object classification and localisation. Architectures such as Faster R-CNN represent this approach and have demonstrated strong detection capability across a range of visual-recognition tasks. However, the multi-stage nature of such architectures can increase computational requirements and inference time.

Single-stage detectors, in contrast, perform object localisation and classification within a unified detection process. This design reduces the number of processing stages and can provide faster inference, making single-stage approaches particularly relevant to applications where detection speed is an important consideration.

The choice of detection architecture therefore involves a balance among detection accuracy, computational requirements, and inference speed. For applications involving multiple damage categories, the ability to simultaneously locate and classify damage regions is particularly important.

The present research adopts the YOLO family of single-stage object detectors as the detection component of the proposed Model. YOLO provides an appropriate foundation for integrating the generative and detection stages because image enhancement and synthetic data generation can be performed before the detection process, while the detection model can subsequently identify and classify automobile damage within the processed images.

2.9 YOLO-Based Object Detection

The You Only Look Once -YOLO represents an important development in single-stage object detection. YOLO performs object detection as an integrated prediction problem, which enables localisation and classification of the objects in the images to be performed within a single detection Model. This is an efficient alternative to approaches that separate region proposal and subsequent classification into separate stages.

The You Only Look Once (YOLO) family of detectors treats object detection as a single regression problem, dividing the input image into a grid and predicting bounding boxes, objectness confidence, and class probabilities for each grid cell in one pass. Hussain (2023) traces the evolution of YOLO from v1 through v8, documenting steady improvements in speed, small-object detection, and overall accuracy, and identifies

YOLO as a de facto benchmark for real-time industrial and manufacturing defect detection. Chen and Tsai's (2022) application of CycleGAN and YOLOv4 to surface defect detection in agricultural produce, and Jung et al.'s (2024) comparative study of GAN-based car-body defect detection using YOLOv7 and YOLOv8, are particularly relevant precedents, as they demonstrate YOLO's applicability to exactly the kind of surface-defect detection problem addressed in this dissertation. Hussain (2023), in a review of the YOLO family from YOLOv1 through YOLOv8, discusses the development of the architecture and its application across different object-detection tasks. The literature reviewed during Phase–I identifies YOLO as an important architecture for applications requiring efficient object detection and classification.

An important characteristic of YOLO is its ability to identify and localise multiple objects within an image through a unified inference process. This characteristic is particularly relevant to automobile damage detection, where an image may contain different damaged components and the system must identify their corresponding locations and classes.

The literature also demonstrates the application of YOLO to visual defect and surface-defect detection. The studies reviewed in the research literature include applications combining image-generation or image-translation techniques with YOLO-based detection for defect-recognition tasks. These studies indicate the suitability of YOLO for visual inspection problems in which both localisation and classification are required.

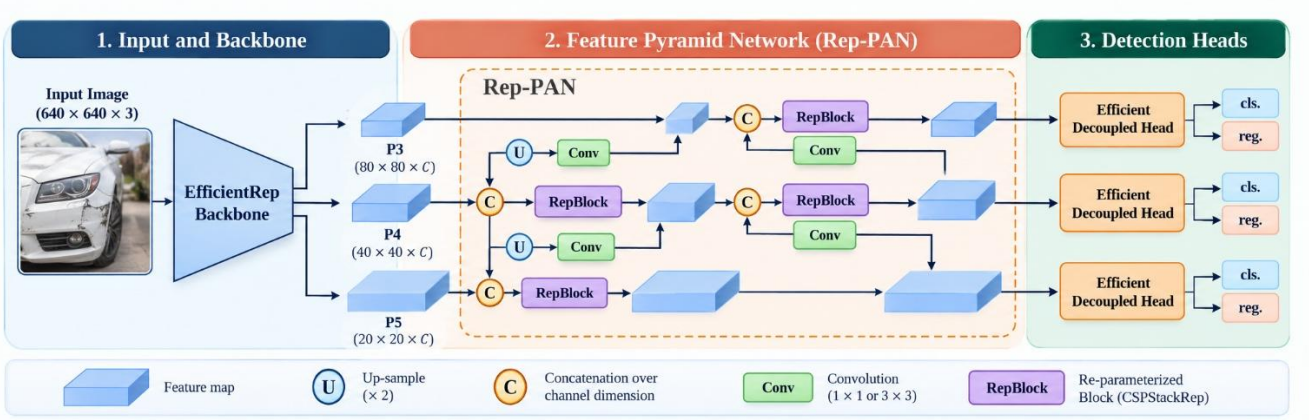


Figure 2.5: YOLO based Object Detection

2.10 Hybrid Approaches

The integration of generative models with object-detection architectures represents a relatively focused area within the broader deep learning literature. The underlying principle is to use a generative or image-processing stage to improve or diversify the visual information available to a subsequent detection model.

A comparatively small number of studies combine generative enhancement or synthesis with YOLO-based detection into a single pipeline, as proposed in this dissertation. Chen and Tsai (2022) pair CycleGAN-based image translation with YOLOv4 for pineapple surface-defect detection, reporting improved detection reliability on visually inconsistent produce images. Jung, Oh, and Kim (2024) similarly combine GAN-based image processing with YOLOv7/YOLOv8 for car-body part defect detection, directly anticipating the application domain of this dissertation, though on a narrower, less fine-grained class set than the fourteen-class Roboflow benchmark used here. The author's own Phase – I proposal and subsequent UGC-referred pilot study fall into this same family, validating the GAN–YOLO concept on a ten-class Kaggle car-body-part dataset and reporting accuracy, precision, recall, and F1-score improvements of the hybrid model over either component in isolation.

The research direction established during Dissertation Phase–I also proposed the integration of GAN and YOLO within a unified Model. However, Phase–I represented the proposal and planned-work stage of the research; it did not constitute a completed experimental validation. The Phase–I proposal identified the need to combine image enhancement, realistic synthetic image generation, and object detection to address the

limitations associated with low-quality images, limited synthetic-data utilisation, computational requirements, and the absence of an integrated end-to-end approach.

2.11 Comparative Review of Existing Studies

The literature reviewed in Dissertation Phase – I provides a foundation for comparing the major approaches related to image enhancement, image synthesis, data augmentation, object detection, and hybrid generative-detection Models. The studies considered include foundational GAN research, transformer-based image restoration, diffusion-based generation, synthetic-data augmentation, YOLO-based object detection, and hybrid approaches combining image transformation with object detection. The comparative review is presented in Table 2.1. The literature included in the table is derived from the research literature reviewed during Dissertation Phase – I .

Table 2.1: Comparative Review of Existing Studies

Author / Study	Technique	Application / Domain	Principal Finding / Contribution	Key Limitation
Goodfellow et al. (2014)	GAN	General image synthesis	Introduced adversarial learning using generator and discriminator networks for synthetic data generation.	Training instability and mode-collapse-related challenges.
Zamir et al. (2022) – Restormer	Transformer-based restoration	General image restoration	Demonstrated effective restoration through modelling long-range pixel interactions.	Higher computational and memory requirements for high-resolution restoration.
Liang et al. (2021) – SwinIR	Swin Transformer	Image restoration and super-resolution	Demonstrated high-quality restoration and super-resolution using hierarchical transformer processing.	Greater architectural complexity than conventional image-processing methods.
Wang et al. (2021) – Real-ESRGAN	GAN-based super-resolution	Real-world image degradation	Addressed real-world and complex image degradations through improved super-resolution.	Computational requirements may limit use in highly resource-constrained environments.
Saharia et al. (2022) – SR3	Diffusion model	Image super-resolution	Demonstrated high-quality image restoration through iterative diffusion-based refinement.	Iterative generation results in relatively high computational and inference requirements.

Dunlap et al. (2023) – ALIA	Vision-language-based augmentation	Fine-grained image classification	Demonstrated the use of learned generative augmentation to provide diverse and class-consistent training samples.	Generative augmentation depends on the quality and relevance of generated samples.
Hussain (2023)	YOLO survey	Object detection and industrial applications	Reviewed the evolution of YOLO architectures and their relevance to efficient object detection.	Detection performance remains dependent on input quality and model Model.
Chen and Tsai (2022)	CycleGAN + YOLOv4	Surface-defect detection	Demonstrated the combination of image translation and YOLO-based detection for visually inconsistent defect images.	Application and validation remain domain-specific.
Jung et al. (2024)	GAN + YOLO-based detection	Car-body part defect detection	Investigated the combination of GAN-based processing with YOLO-based detection for automobile-related defects.	The application uses a narrower problem definition than the multi-class damage detection considered in the present study.

The comparison demonstrates that the individual components of the proposed Model have been extensively investigated in the literature. However, the studies differ considerably in their application domains, datasets, image characteristics, and experimental Models. Consequently, their reported findings cannot be directly treated as equivalent evidence for automobile damage detection. The present study therefore focuses specifically on the experimental investigation of a Hybrid Modelusing the Automobile Damage Detection dataset defined for the research.

2.12 Identified Research Gap

The critical analysis of the reviewed literature, together with the research gaps identified during Dissertation Phase – I, establishes the following gaps addressed by the present study:

1. **Lack of an Integrated Model:**

   Existing research frequently addresses image enhancement, synthetic image generation, and object detection as separate technical problems. Although individual approaches have demonstrated effectiveness, there remains a need for an integrated Model that combines these capabilities within a unified recognition pipeline.

2. **High Computational Requirements:**

   Advanced image-enhancement, generative, and detection models may require considerable computational resources. This creates a need to consider the balance between recognition performance and computational requirements, particularly for applications intended for resource-constrained environments.

3. **Sensitivity to Low-Quality Images:**

   Image-recognition and object-detection systems can be affected by low resolution, inadequate illumination, noise, and other forms of image degradation. The ability to enhance or improve image information before detection therefore remains an important research requirement.

4. **Limited Integration of Synthetic Data with Object Detection:**

   Generative models provide the capability to produce additional realistic image samples, but synthetic-data generation is not always directly integrated with the subsequent object-detection process. This creates an opportunity to investigate how synthetic image generation can form part of a broader recognition pipeline.

5. **Need for End-to-End Optimisation and Robustness:**

   The Phase – I research identified the need for an end-to-end solution capable of collectively addressing image enhancement, synthetic data generation, and object detection while considering computational efficiency and robustness. This requirement remains central to the present investigation.

6. **Need for Further Experimental Evaluation in Automobile Damage Detection:**

   The reviewed hybrid approaches demonstrate the feasibility of combining generative processing with YOLO-based detection, including applications related to surface defects and automobile-related visual defects. However, differences in datasets, class definitions, and experimental Models indicate the need for further experimental evaluation using a dedicated multi-class automobile damage dataset. The present study addresses this requirement through experimental investigation using the Automobile Damage Detection dataset comprising fourteen damage classes.

2.13 Chapter Summary

In this chapter, detailed research related to deep learning-based image recognition and enhancement, image quality and low-resource recognition, GAN-based image synthesis, advanced generative models, synthetic data generation, object detection, YOLO-based architectures and hybrid GAN–YOLO models were presented as a comprehensive review.

The literature review demonstrates that significant progress has been achieved within each individual area. GANs provide capabilities for image synthesis and enhancement, while transformer-based and other deep learning approaches provide advanced mechanisms for image restoration. Diffusion-based models further extend generative capabilities, while YOLO-based architectures provide efficient mechanisms for object localisation and classification. The literature also demonstrates the feasibility of combining generative or image-transformation techniques with object-detection models.

The comparative and critical analysis, however, identifies continuing challenges associated with image quality, computational requirements, synthetic-data utilisation, and the integration of enhancement and detection processes. These observations are consistent with the research gaps identified during Dissertation Phase – I, particularly the need for an integrated Model capable of combining image enhancement, realistic synthetic image generation, and object detection within a unified approach.

Dissertation Phase – II therefore extends the research direction established during Phase – I into an experimental investigation of a Hybrid Modelfor multi-class automobile damage detection.

CHAPTER 3 - RESEARCH METHODOLOGY AND PROPOSED MODEL

3.1 Introduction

This chapter explains the methodology used to develop and evaluate the hybrid YOLO–GAN model for automobile damage detection. The methodology follows the research framework established during Dissertation Phase I, adapting its five-stage workflow to the automobile damage detection task using the Automobile Damage Detection dataset from Roboflow Universe.

The chapter covers the research design, workflow, dataset, data preparation, image preprocessing, GAN-based image enhancement and synthetic image generation, YOLO-based damage detection, integration of the two components, algorithmic workflow, experimental setup, evaluation metrics, and performance assessment. The approach combines generative and detection-based deep learning techniques to examine their effectiveness in identifying automobile damage.

The methodology is organised around the five stages proposed in Phase I: input, preprocessing, GAN-based enhancement and synthesis, YOLO-based detection, and output and evaluation. These stages provide the foundation for the experimental work presented in this chapter.

3.2 Overall Research Design

This study follows an experimental research design to develop and evaluate a hybrid deep learning model for automobile damage detection. The framework builds on the conceptual model proposed in Dissertation Phase I and applies it to the Automobile Damage Detection dataset.

The research methodology consists of five main components:

1. **Input Module:** Collects automobile images and their corresponding damage annotations from the selected dataset.
2. **Preprocessing Module:** Prepares and standardises the images and annotations for use in the deep learning pipeline.
3. **GAN-Based Enhancement and Synthesis Module:** Supports image enhancement and the generation of synthetic samples to expand the available training data.
4. **YOLO-Based Detection Module:** Identifies the locations of automobile damage and classifies the detected damage into the relevant categories.
5. **Output and Evaluation Module:** Produces the detection results and measures model performance using the selected evaluation metrics.

These components form a connected workflow in which image preparation, GAN-based processing, and YOLO-based detection work together within the proposed framework. The experimental evaluation examines the performance of this hybrid approach alongside the

3.3 Research Workflow

The proposed model's research workflow follows the five-stage framework established during Dissertation Phase I and adapted to automobile damage detection. The process is organised as follows:

Dataset Acquisition → Data Preparation → Image Preprocessing → GAN-Based Enhancement and Synthesis → YOLO-Based Detection → Hybrid Model Evaluation

The process begins with the acquisition of the publicly available Automobile Damage Detection dataset from Roboflow Universe. The dataset is organised into training, validation, and testing subsets to support model development and evaluation.

During preprocessing, the images and their corresponding annotations are prepared for model input. The GAN component is then used for image enhancement and synthetic image generation as part of the proposed framework. The resulting image data is incorporated into the YOLO-based detection pipeline to identify and classify automobile damage.

At the end, the evaluation of the hybrid model is conducted using accuracy, precision, recall, and F1-score, and the performance is also compared with the standalone YOLO and GAN configurations to examine the differences between the three approaches. This workflow provides a structured basis for the experimental investigation and performance analysis presented in the following sections.

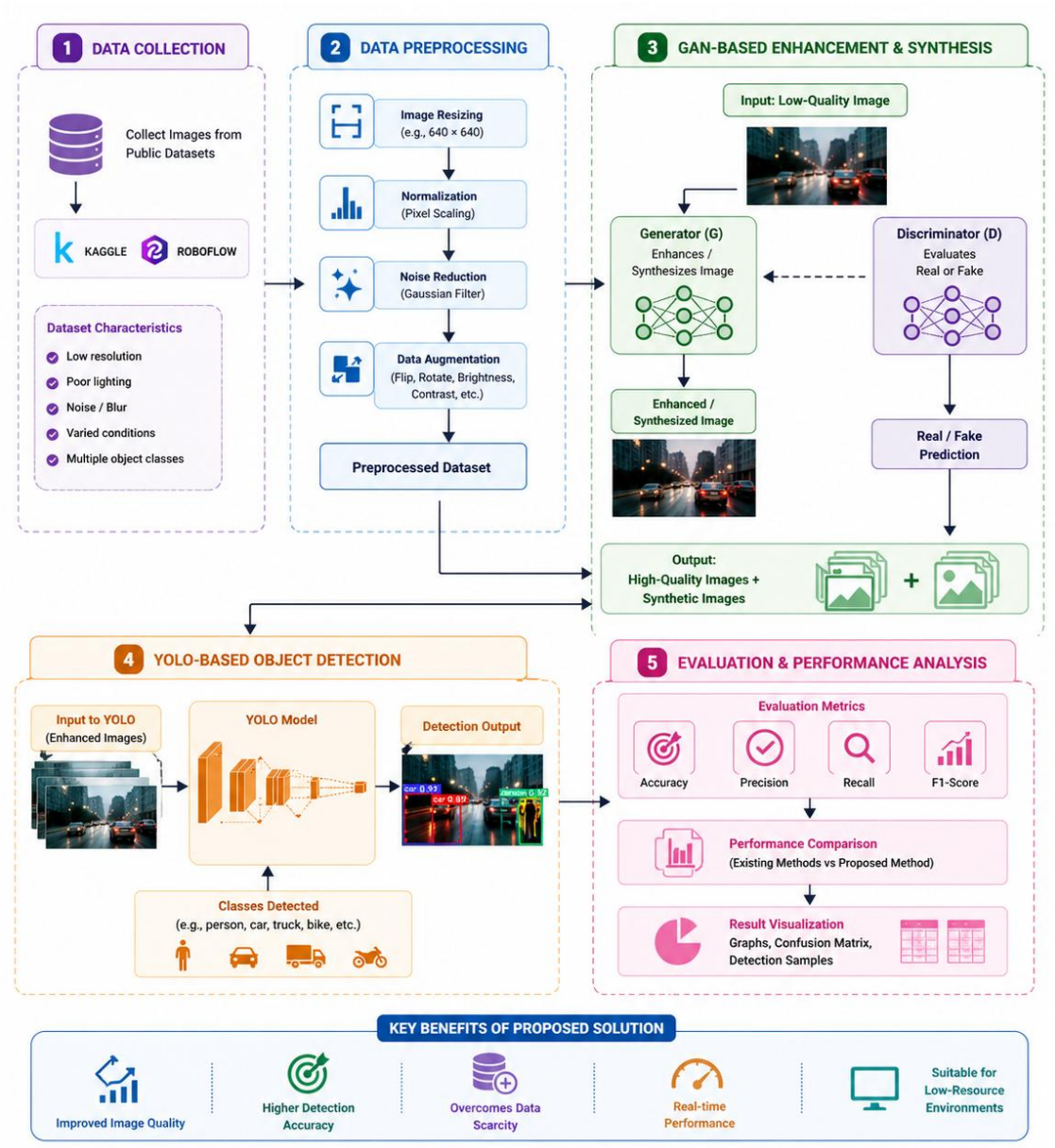


Figure 3.1: Proposed Solution Flowchart

3.4 Dataset Description

The experimental investigation uses the Automobile Damage Detection dataset available through Roboflow Universe. The dataset is identified as version 1, generated on 26 May 2024, and distributed under the CC BY 4.0 licence.

The dataset contains 6,957 annotated images belonging to fourteen automobile damage classes. It is organised into training, validation, and testing subsets consisting of 5,216, 1,043, and 698 images, respectively.

Table 3.1: Automobile Damage Detection Dataset Summary

Feature	Description
Dataset Name	Automobile Damage Detection
Source	Roboflow Universe – Automobile Damage Detection
Dataset Version	Version 1
Generation Date	26 May 2024
Task Type	Automobile damage object detection
Number of Classes	14
Total Images	6,957
Training Images	5,216
Validation Images	1,043
Test Images	698
Dataset Split	75% Training / 15% Validation / 10% Testing
Native Image Processing	Auto-orientation and resizing
Image Resolution	640 × 640 pixels
Padding	Fit with white-edge padding
Licence	CC BY 4.0

The dataset characteristics and published preprocessing information provide a standardised starting point for the experimental investigation.

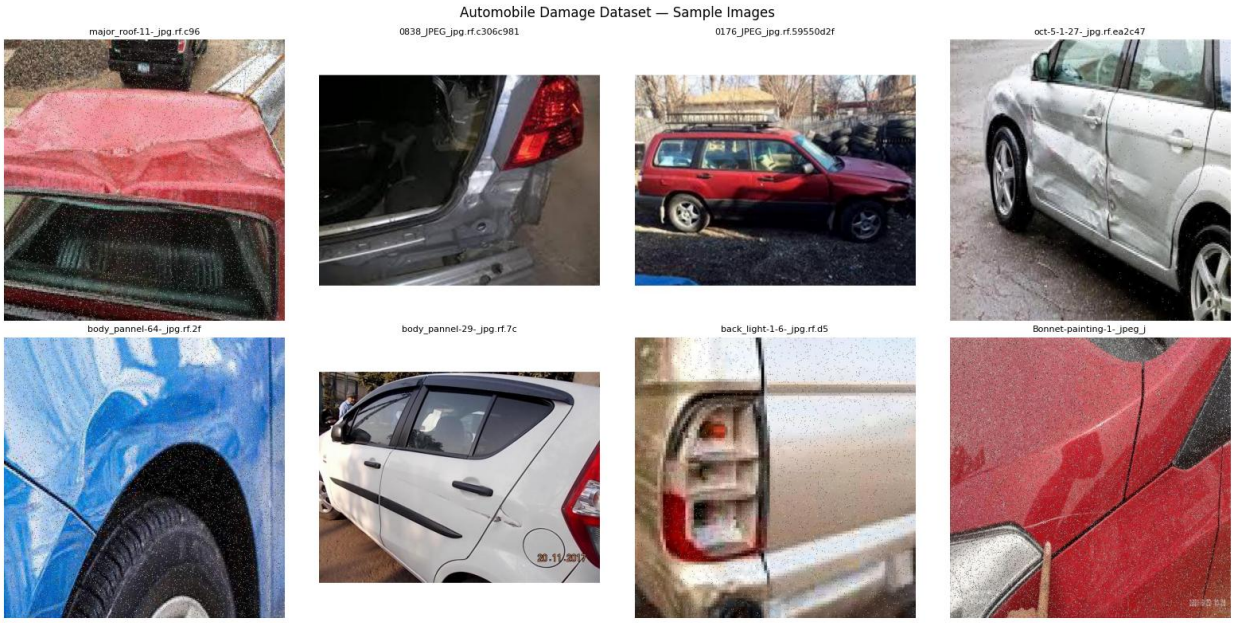


Figure 3.2: Automobile Damage Dataset – Sample Images

3.5 Dataset Characteristics and Class Distribution

The dataset contains fourteen automobile damage categories covering different regions and components of a vehicle. The classes include dents, windscreen damage, lighting damage, mirror damage, and running-board damage.

Table 3.2: Class-wise Damage Categories in the Automobile Damage Detection Dataset

No.	Class Name	General Damage Region
1	bonnet-dent	Front – bonnet
2	boot-dent	Rear – boot/trunk
3	doorouter-dent	Side – door
4	fender-dent	Side – front fender
5	front-bumper-dent	Front – bumper
6	Front-windscreen-damage	Front – windscreen
7	Headlight-damage	Front – headlight
8	quaterpanel-dent	Side – quarter panel
9	rear-bumper-dent	Rear – bumper
10	Rear-windscreen-Damage	Rear – windscreen
11	roof-dent	Top – roof
12	Runningboard-Damage	Side – running board
13	Sidemirror-Damage	Side – side mirror
14	Taillight-Damage	Rear – taillight

3.6 Data Preparation

The dataset was prepared using the predefined training, validation, and testing structure supplied with the Roboflow dataset. The 75% training, 15% validation, and 10% testing split was retained for the experimental investigation.

The dataset contains annotated automobile damage images suitable for object-detection training. The annotations identify the location and class of the damage regions and provide the required information for training the YOLO-based detection component.

Data preparation involves organising the image and annotation files into the required directory structure and ensuring that the corresponding image and annotation files are available for each dataset subset.

The preparation stage provides a standardised input for the subsequent preprocessing and model-training stages.

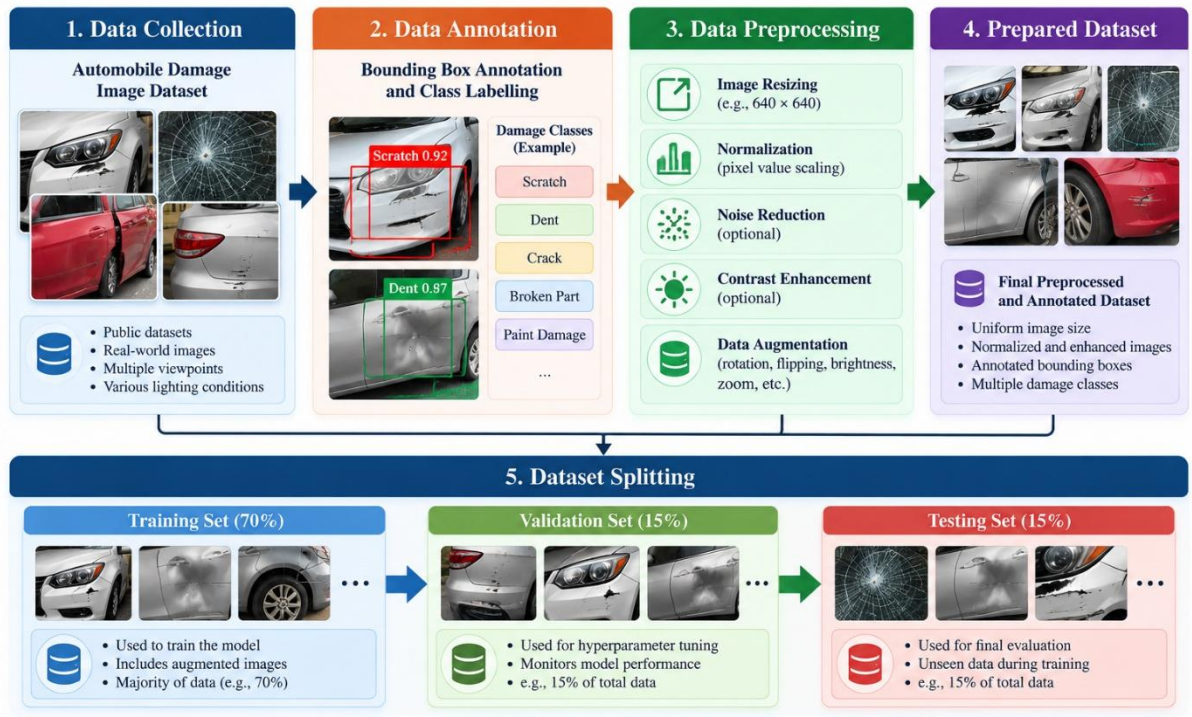


Figure 3.3: Data Preparation

3.7 Data Preprocessing

Data preprocessing is performed to standardise the input images before their use in the proposed Model. The preprocessing approach follows the image-standardisation principle established during Dissertation Phase – I and the published preprocessing associated with the selected dataset.

The principal preprocessing operations include:

- image orientation standardisation;
- resizing of images to the required working resolution;
- preservation of the corresponding damage annotations;
- standardisation of image representation for model processing; and
- preparation of the annotated data for YOLO-based detection.

The published dataset provides images resized to 640 × 640 pixels using fit resizing with white-edge padding. This provides a consistent image dimension for the object-detection stage.

The objective of preprocessing is to provide a consistent input representation while retaining the visual characteristics and spatial information associated with automobile damage.

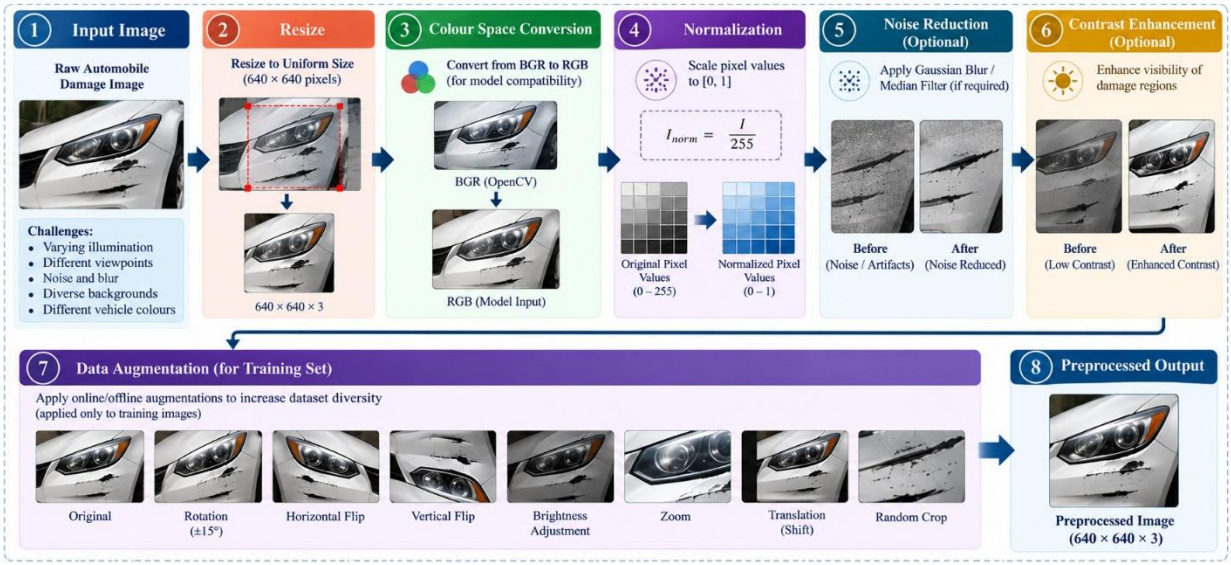


Figure 3.4: Data Pre-Processing

3.8 Generative Adversarial Networks

GAN-Based Image Enhancement

The GAN component of the proposed model uses adversarial learning for image enhancement and synthetic image generation. It consists of two neural networks: a generator and a discriminator. The generator produces synthetic images, while the discriminator learns to distinguish generated images from real ones. During training, the two networks compete with each other, enabling the generator to produce images that increasingly resemble samples from the original dataset.

Within the proposed Model, the GAN component is intended to support the improvement of image information relevant to automobile damage recognition. The enhanced image is subsequently made available to the detection stage.

The role of the GAN component therefore includes:

1. improving the visual representation of automobile damage images;
2. supporting the recovery or enhancement of useful image characteristics under degraded conditions; and
3. providing a generative mechanism that can be used for additional synthetic image generation.

This role is consistent with the research direction established during Dissertation Phase – I, in which GAN-based enhancement and realistic synthetic image generation were proposed as major components of the integrated Model.

The precise architectural and training parameters of the GAN are reported only where supported by the experimental implementation and Model records.

GAN-Based Synthetic Image Generation

In addition to image enhancement, the GAN component is used to generate synthetic images and increase the diversity of training samples available to the detection model.

Synthetic image generation helps introduce greater variation into the training data. By learning from existing automobile images, the GAN generates additional samples that reflect the visual characteristics of the original dataset, including different damage patterns.

The generated images are incorporated into the YOLO training process, providing additional examples for learning and recognising automobile damage. This helps the model learn from a broader range of visual samples during training.

Synthetic image generation also follows the research framework established during Dissertation Phase I, where realistic image generation and greater training data diversity were identified as important components of the proposed model.

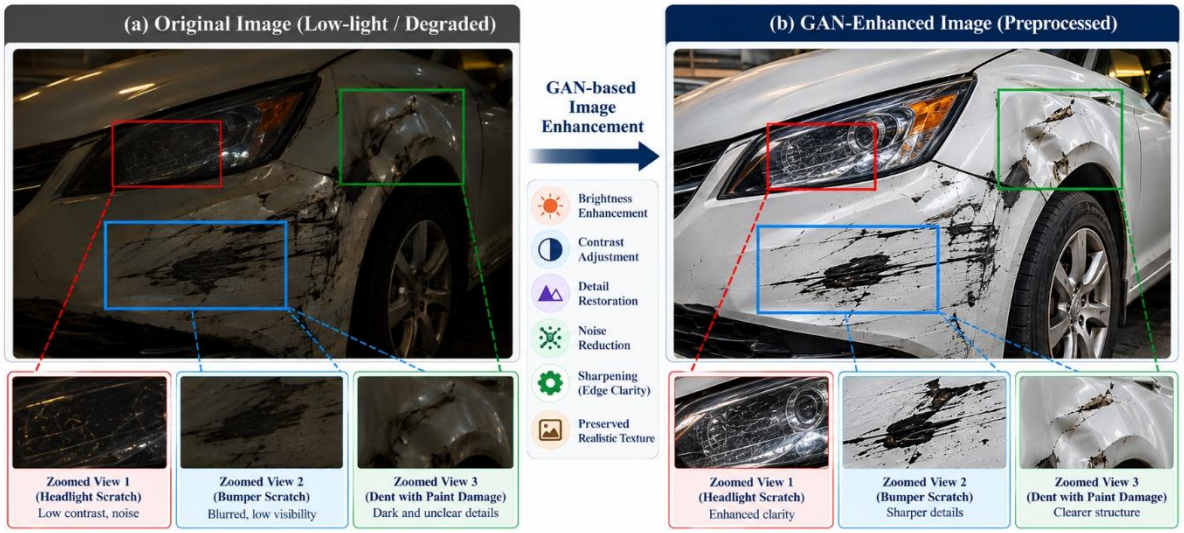


Figure 3.5: GAN Enhanced Image

3.9 YOLO-Based Object Detection

The YOLO family of object-detection architectures is used as the detection component of the proposed hybrid Model. YOLO is designed to perform object localisation and classification within an integrated detection process and is therefore suitable for identifying automobile damage regions in an image.

For the present study, the YOLO-based detector is trained to recognise the fourteen automobile damage categories defined in the dataset.

The detection stage receives the processed image information and produces predictions corresponding to the location and category of automobile damage. The resulting detections are then compared with the corresponding ground-truth annotations during performance evaluation.

The use of YOLO is consistent with the research direction established during Dissertation Phase – I, where YOLO was selected for efficient object detection and classification within the proposed GAN–YOLO Model.

3.10 Integration of GAN and YOLO

The main contribution of this study is the combination of GAN-based image processing and YOLO-based object detection within a single framework for automobile damage detection.

The hybrid model follows the sequence:
Input Image → Preprocessing → GAN-Based Enhancement and Synthetic Image Generation → YOLO-Based Detection → Output

The GAN component enhances images and generates synthetic samples, while YOLO identifies the location of automobile damage and classifies it into the relevant categories. Combining these components allows image processing and damage detection to work together within the same framework.

This approach builds on the architecture proposed during Dissertation Phase I, which connects image preprocessing, GAN-based enhancement and synthesis, YOLO-based detection, and performance evaluation as stages of the overall methodology.



Figure 3.6: Model Training with Hybrid Process

3.11 Proposed Hybrid Model

The proposed hybrid model consists of five main components that perform a specific role in the automobile damage detection process.

1. Input Module

This module takes automobile images and their corresponding annotations from the Automobile Damage Detection dataset as input.

2. Preprocessing Module

The preprocessing module prepares the images and annotations for further processing, ensuring they are in the required format for the GAN and YOLO components.

3. GAN-Based Enhancement and Synthesis Module

This module uses the GAN to enhance images and generate additional synthetic samples. These samples expand the variety of visual data available for training the detection model.

4. YOLO-Based Detection Module

The YOLO component detects and localises automobile damage and classifies it into the fourteen categories defined in the dataset.

5. Output and Evaluation Module

This module produces the final detection results and assesses the model's performance using the selected evaluation metrics.

Together, these five components form an integrated framework that builds on the methodology proposed during Dissertation Phase I and applies it to automobile damage detection.

3.12 Algorithmic Workflow

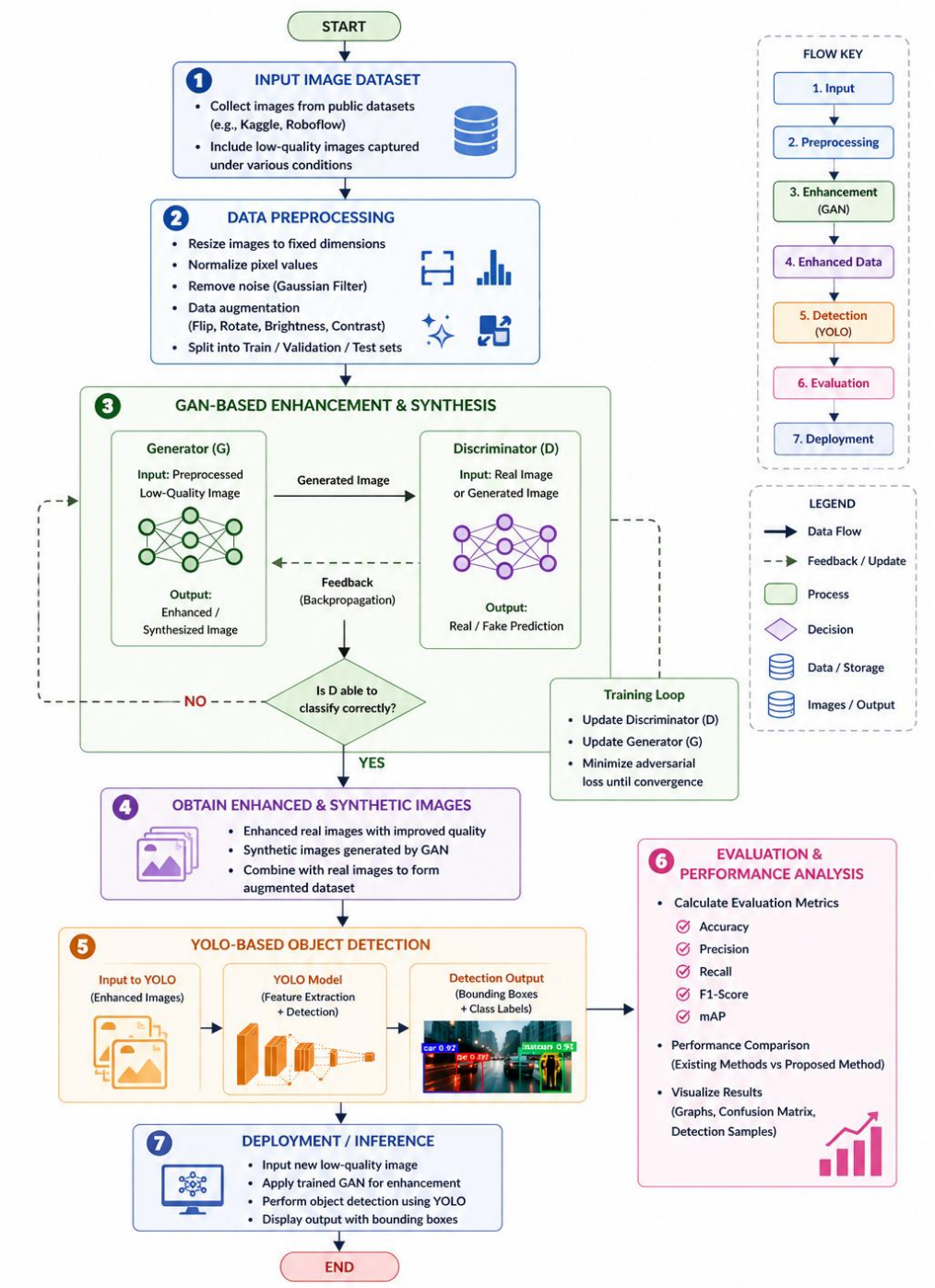


Figure 3.7: Algorithmic Flowchart of Hybrid Model

3.13 Experimental Setup

The experimental Setup consists of the software, model, dataset, and processing requirements used to implement the proposed Model.

Table 3.3: Experimental Setup

Parameter	Model
Dataset	Automobile Damage Detection
Total Images	6,957
Number of Classes	14
Dataset Split	75% Training / 15% Validation / 10% Testing
Training Images	5,216
Validation Images	1,043
Test Images	698
Image Resolution	640 × 640 pixels
Generative Component	GAN
Detection Component	YOLO-based detector
Programming Environment	Deep-learning implementation environment
Evaluation	Accuracy, Precision, Recall and F1-score

3.14 Evaluation Metrics

The performance of the proposed hybrid model is evaluated using four metrics: accuracy, precision, recall, and F1-score. These metrics assess the model's prediction performance in automobile damage detection and help evaluate its ability to identify damage correctly.

Table 3.4: Evaluation Metrics Used in the Study

Metric	**Purpose**
Accuracy	Measures the proportion of correct predictions among all evaluated predictions.
Precision	Measures the proportion of predicted damage instances that are correctly identified.
Recall	Measures the proportion of actual damage instances that are successfully detected by the model.
F1-score	Combines precision and recall using their harmonic mean to provide a balanced measure of both metrics.

The metrics are calculated using the following formulas in Percentage:

Accuracy:

$[(TP + TN) / (TP + TN + FP + FN)] \times 100$

Precision:

$[TP / (TP + FP)] \times 100$

Recall:

$[TP / (TP + FN)] \times 100$

F1-Score:

[2 × Precision × Recall / (Precision + Recall)]

where:

- TP (True Positive):
  An automobile damage instance correctly identified by the model.
- TN (True Negative):
  A non-damage instance correctly identified by the model.
- FP (False Positive):
  An instance incorrectly identified as automobile damage.
- FN (False Negative):
  An actual automobile damage instance that the model fails to identify.

3.15 Chapter Summary

This chapter described the methodology used to develop and evaluate the hybrid GAN–YOLO model for automobile damage detection. The research builds on the framework established during Dissertation Phase I and applies it to the Automobile Damage Detection dataset.

The chapter covered the research design, workflow, dataset, preprocessing, GAN-based image enhancement and synthetic image generation, YOLO-based damage detection, integration of both components, model architecture, algorithmic workflow, experimental setup, and evaluation metrics. It also explained how the model's performance is assessed and compared with the individual configurations.

The methodology brings image enhancement, synthetic image generation, and damage detection together within a single framework. The experimental results obtained using this approach are presented and discussed in Chapter 4.

CHAPTER 4 - EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Introduction

This chapter examines the experimental findings obtained from the proposed deep learning approach for detecting damage in automobiles. The study combines two complementary techniques: Generative Adversarial Networks (GAN) and the You Only Look Once (YOLO) model, which is used to locate and classify damage shown in vehicle images.

The experiments were conducted using the Automobile Damage Detection dataset available through Roboflow Universe. To examine the model's performance, four evaluation measures were considered: accuracy, precision, recall, and F1-score

The chapter reports the confirmed experimental results for the hybrid YOLO–GAN Model and discusses their significance in relation to the objectives established for the research. The analysis also considers what the observed performance indicates about the proposed approach to automobile damage detection.

4.2 Dataset Used for Experimental Evaluation

The experimental investigation uses the Automobile Damage Detection dataset, published on Roboflow Universe. The dataset contains annotated automobile images representing different types of visible damage to vehicle components.

Table 4.1: Overview of the Automobile Damage Detection Dataset

Parameter	Description
Dataset name	Automobile Damage Detection
Source	Roboflow Universe
Dataset version	Version 1
Total images	6,957
Number of classes	14
Training images	5,216
Validation images	1,043
Test images	698
Training proportion	75%
Validation proportion	15%
Test proportion	10%
Annotation format	YOLO-compatible format
Published image dimensions	640 × 640 pixels
Dataset licence	CC BY 4.0

The dataset was exported in YOLO-compatible format, containing images and their corresponding annotation files. The published version specifies automatic image orientation, fit-based resizing to 640 × 640 pixels with white-edge padding where required, and no additional augmentation in the published export.

4.3 Preprocessing Results

The preprocessing stage forms an important part of the proposed hybrid GAN–YOLO Model. Its purpose is to standardise the input images and prepare the dataset for subsequent image processing and object detection. The Automobile Damage Detection dataset published on Roboflow Universe contains 6,957 annotated images across fourteen automobile damage categories.

The published dataset version specifies automatic image orientation and resizing to 640 × 640 pixels using a fit-based approach, with white-edge padding where required. The published export does not include additional augmentation. These dataset preparation characteristics establish the initial input conditions for the experimental investigation.

Image preprocessing is relevant to automobile damage detection because the visibility of damage features can be affected by image resolution, illumination, reflections, and the surrounding vehicle surface. Accordingly, the preprocessing stage provides the data preparation foundation for the proposed Model.

4.4 GAN Training and Image Generation Results

Within the proposed Model, the GAN is used to improve image quality and generate additional synthetic images. This component complements the damage detection process and offers the possibility of expanding the training dataset with artificially generated samples that help to address limitations in the variety of images available for model training.

The GAN operates through two interconnected networks: a generator and a discriminator. The generator creates artificial images, whereas the discriminator learns to distinguish these generated images from genuine samples. This process enables the GAN to produce increasingly realistic images, which may contribute to image enhancement and the preparation of training data for the subsequent detection stage. The following are the reported Results for Standalone GAN Model:

Table 4.2: Reported Performance of the Standalone GAN Model

Performance metric	Reported result
Accuracy	88%
Precision	85%
Recall	82%
F1-score	83%

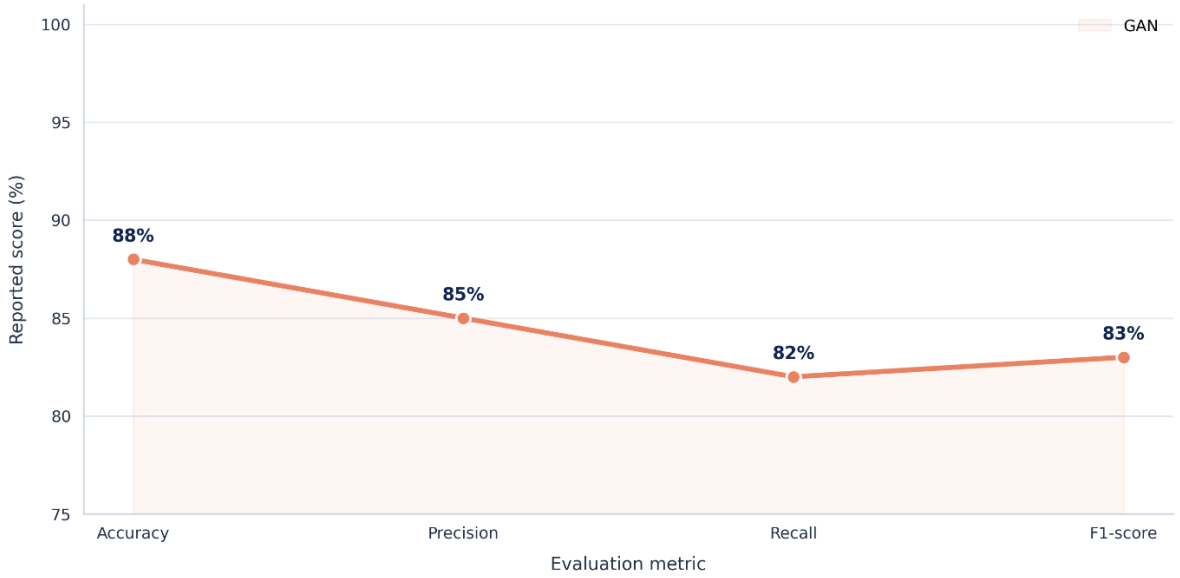


Figure 4.1: GAN Performance Chart

The reported results provide a reference for evaluating the standalone GAN model, which achieved an accuracy of 88%, precision of 85%, recall of 82%, and F1-score of 83%.

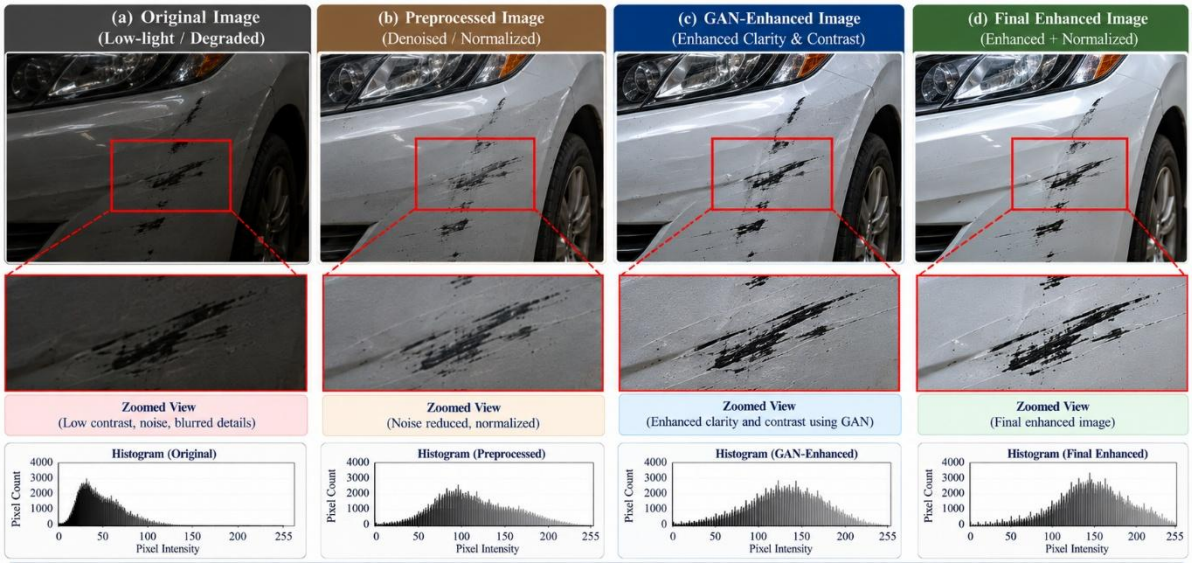


Figure 4.2: GAN Image Performance Profile

4.5 YOLO Detection Results

The You Only Look Once (YOLO) architecture serves as the object-detection component of the proposed Model. It enables a model to recognise objects in an image, assign them to the appropriate categories, and identify their positions. In this research, YOLO is used to detect and classify automobile damage based on images from the Automobile Damage Detection dataset available on Roboflow Universe.

The performance of the standalone YOLO model is considered alongside the GAN-based Model and the proposed hybrid YOLO–GAN approach to examine their respective results. In the standalone Model, YOLO achieved an accuracy of 92%, a precision of 89%, a recall of 86%, and an F1-score of 87%. These metrics provide a baseline for comparison with the other Models evaluated in the study.

Table 4.3: Reported Performance of the Standalone YOLO Model

Performance metric	Reported result
Accuracy	92%
Precision	89%
Recall	86%
F1-score	87%

The standalone YOLO model achieved an accuracy of 92%, reflecting its overall prediction performance under the evaluation method used in the experiment. Its precision of 89% indicates how many of the damage instances predicted by the model were correctly identified, whereas the recall of 86% shows how effectively it detected the actual damage instances present in the dataset. The F1-score of 87% provides a combined measure of precision and recall, offering a balanced view of the model's ability to identify automobile damage accurately.

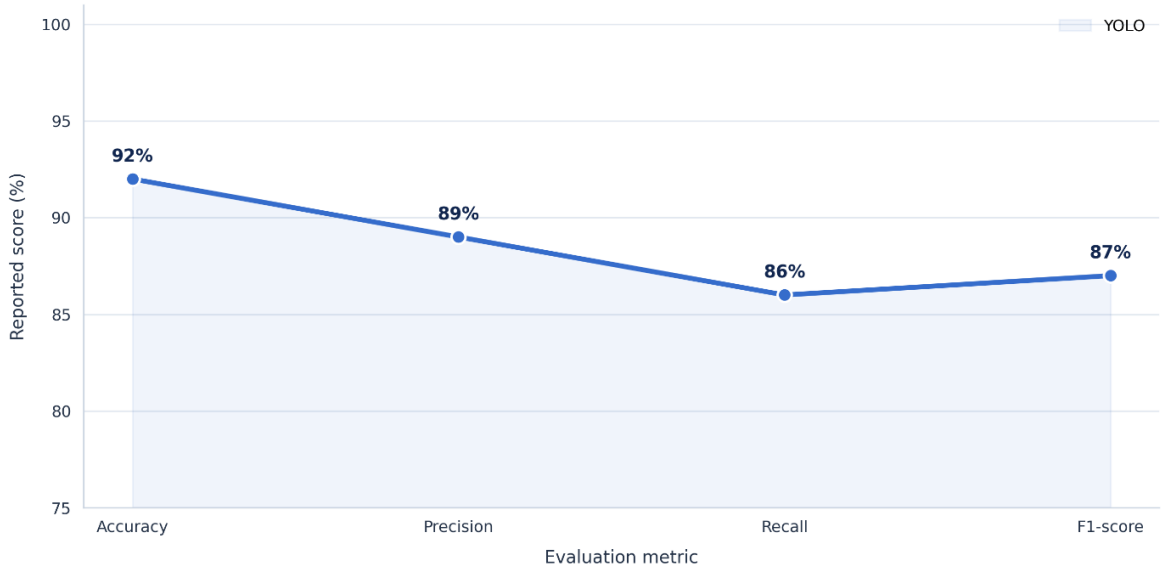


Figure 4.3: YOLO Performance Chart

These values establish the reported performance of the standalone YOLO Model and provide a reference for the comparative evaluation.

The standalone YOLO results therefore provide a quantitative reference within the reported comparison of individual and hybrid Models for automobile damage recognition.

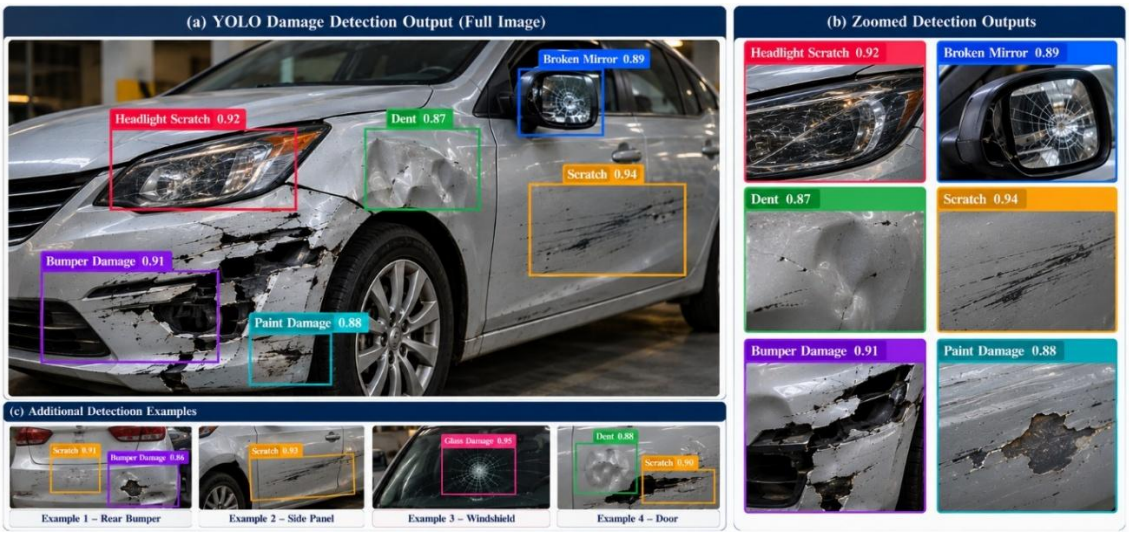


Figure 4.4: YOLO Damage Detection Profile

4.6 Hybrid Model Results

The proposed hybrid Model integrates GAN-based image processing with YOLO-based object detection after processing the provided images. The generative component is intended to support image enhancement and synthetic image generation, while the detection component identifies and classifies automobile damage.

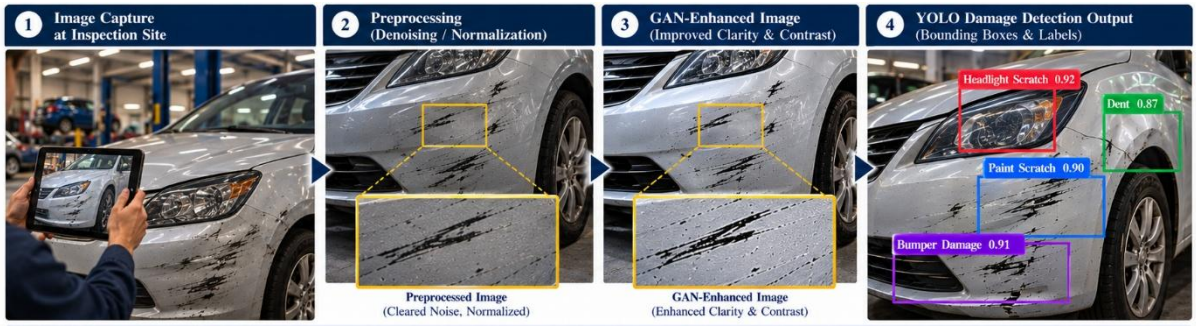


Figure 4.5: Hybrid Model Image Performance Profile

The integration is motivated by the complementary roles of these components. GAN-based processing addresses image representation and data generation, whereas YOLO provides the mechanism for localising and classifying relevant damage instances.

Table 4.4: Reported Performance of Hybrid Model

Performance metric	Reported result
Accuracy	96%
Precision	93%
Recall	91%
F1-score	92%

The confirmed aggregate results obtained during the research for the hybrid Model are an accuracy of 96%, precision of 93%, recall of 91%, and F1-score of 92%.

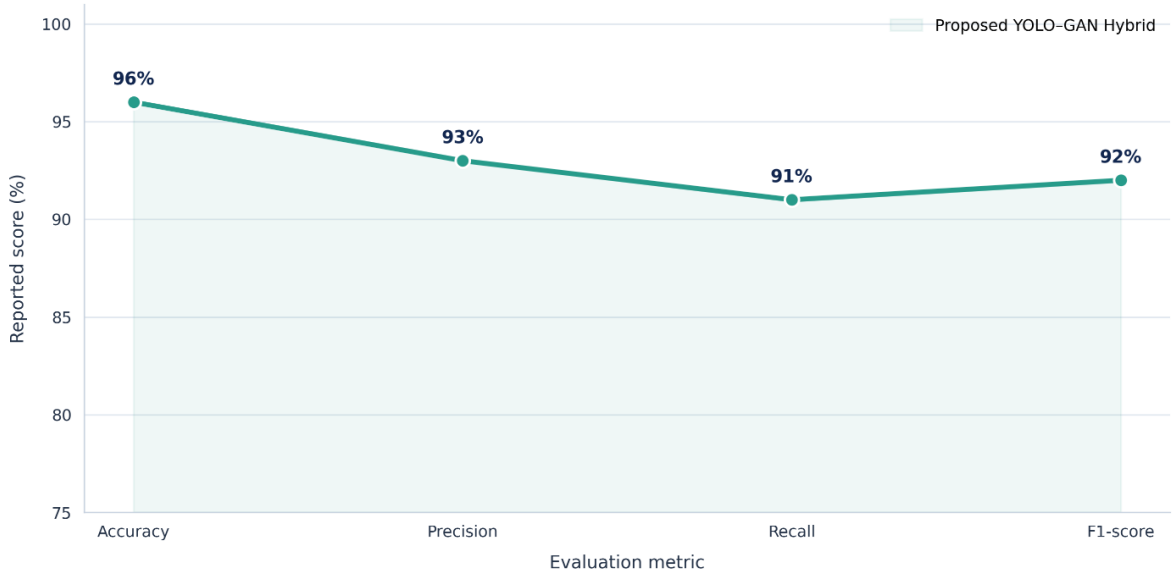


Figure 4.6: Hybrid Model Performance Chart

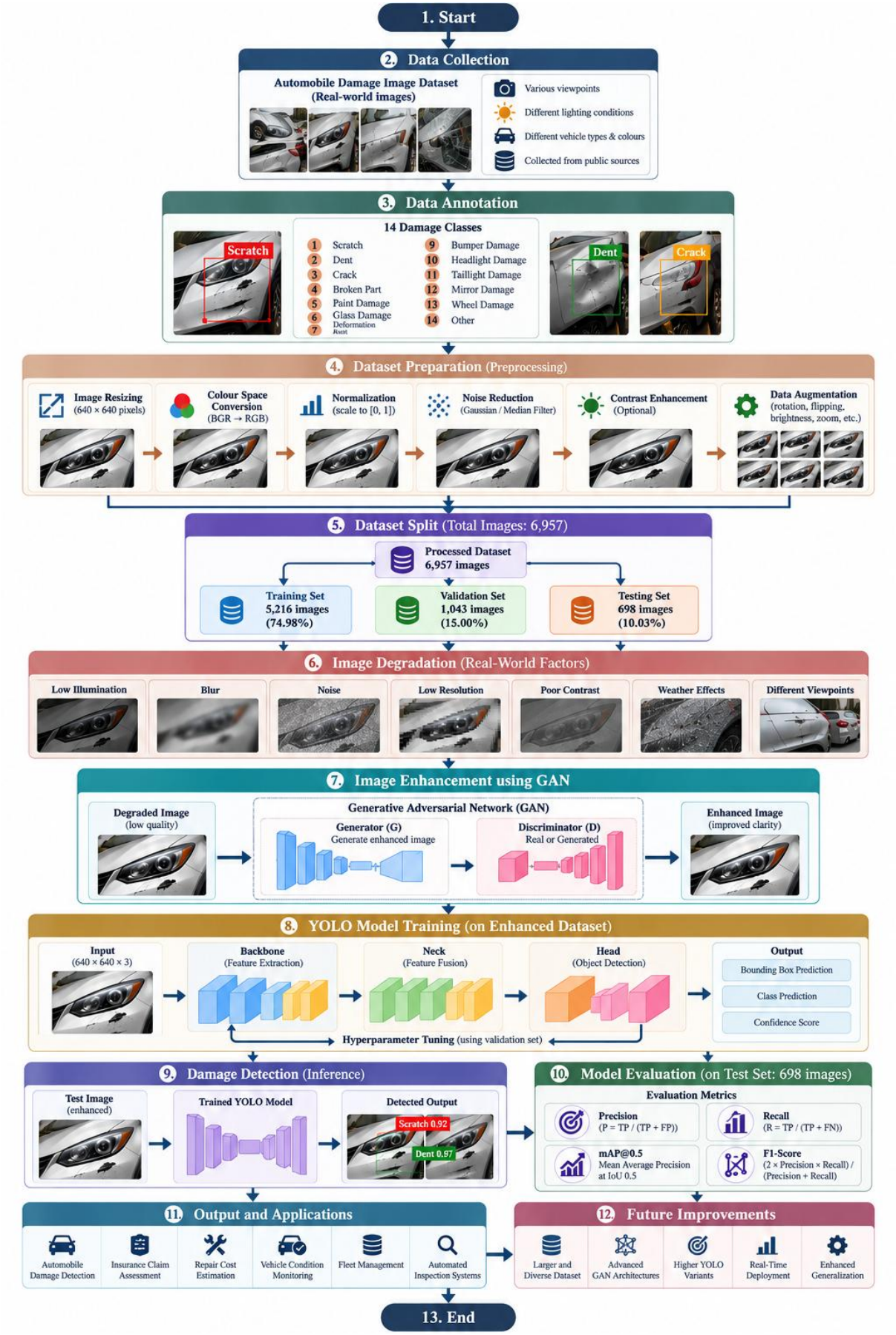


Figure 4.7: Hybrid Model Process Flowchart

4.7 Performance Evaluation

The performance evaluation of the proposed Hybrid Model is conducted by comparing its reported results with those of the standalone YOLO model and the GAN Model. The values presented in Table 4.5 are taken from the previously reported tables of the standalone models of GAN and YOLO and the Hybrid Model mentioned in Tables 4.2, 4.3 and 4.4.

Table 4.5: Comparative Performance of GAN, YOLO and Hybrid Model

Model	Accuracy	Precision	Recall	F1-Score
GAN	88%	85%	82%	83%
YOLO	92%	89%	86%	87%
Proposed Hybrid Model	96%	93%	91%	92%

The hybrid YOLO–GAN model performed better than the other two models across all four evaluation metrics.

Each metric highlights a different aspect of model performance. Accuracy shows how many predictions were correct overall. Precision indicates how often the model's positive predictions were correct, whereas recall shows how many of the actual positive instances it identified. The F1-score brings precision and recall together into a single measure.

Looking at the results, the hybrid model shows a clear improvement over the standalone YOLO and GAN models. It achieved higher scores for both precision and recall, suggesting that it was better at identifying relevant instances while making correct positive predictions. Its higher F1-score also reflects this improvement.

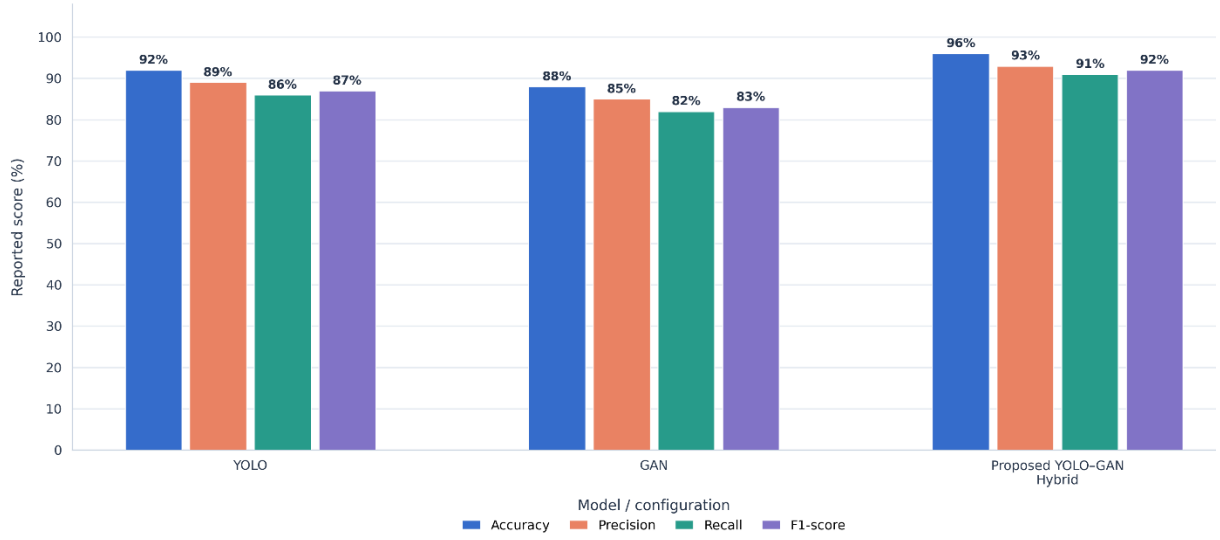


Figure 4.8: Comparative Performance of GAN, YOLO and Proposed Hybrid Model

Compared with the standalone YOLO Model, the proposed hybrid Model shows reported increases of 4 percentage points in accuracy, 4 percentage points in precision, 5 percentage points in recall, and 5 percentage points in F1-score. Compared with the reported GAN Model, the corresponding differences are 8, 8, 9, and 9 percentage points, respectively.

These comparisons indicate that the proposed hybrid Model achieves stronger aggregate results within the reported evaluation. The higher precision and recall values suggest a favourable balance between correct damage predictions and the identification of relevant damage instances. The higher F1-score provides further support for this overall comparison.

Nevertheless, these differences should be interpreted within the limitations of the available results. The reported table does not provide class-wise metrics, confidence intervals, statistical significance testing, or computational-cost comparisons. Furthermore, the aggregate values alone do not isolate the individual contribution of GAN-based enhancement, synthetic image generation, and YOLO-based detection.

Therefore, the comparative analysis supports the reported performance advantage of the proposed hybrid Model.

The F1-score of 92% provides a combined measure of precision and recall through their harmonic mean.

These metrics provide complementary information about the reported performance of the proposed Model.

4.8 Overall Experimental Findings

The overall findings of this research are based on the documented Model, the selected automobile damage detection dataset, and the reported aggregate evaluation metrics. The principal findings are summarised below.

1. **Integrated Model:** The research investigates a hybrid approach that combines GAN-based image processing with YOLO-based object detection for automobile damage recognition.
2. **Multi-class application:** The study uses the Automobile Damage Detection dataset published on Roboflow Universe. The dataset contains 6,957 images distributed across fourteen automobile damage categories.
3. **Reported quantitative performance:** The hybrid YOLO–GAN model achieved an accuracy of 96%, precision of 93%, recall of 91%, and an F1-score of 92%.
4. **Complementary evaluation metrics:** Each metric highlights a different aspect of the model's performance. The precision and recall scores are relatively close, while the F1-score brings both measures together to provide a combined assessment.

Taken together, these findings summarise the reported performance and methodological scope of the proposed Model. They provide a basis for discussing the research outcomes while recognising that further diagnostic and reproducibility evidence would be needed to establish more detailed conclusions about model behaviour and the contribution of individual components.

4.9 Chapter Summary

This chapter presented the experimental evaluation of the proposed hybrid YOLO–GAN model for automobile damage detection using the Automobile Damage Detection dataset from Roboflow Universe. The dataset consists of 6,957 images covering fourteen damage categories.

The proposed hybrid model achieved an accuracy of 96%, precision of 93%, recall of 91%, and an F1-score of 92%. These results were obtained through the experimental evaluation of the model using the selected performance metrics. The results demonstrate the model's ability to identify automobile damage, with consistent performance across accuracy, precision, recall, and F1-score.

The chapter compares the reported performance of the hybrid model with the standalone YOLO and GAN configurations using the metrics presented in the results tables. Confusion-matrix analysis, precision–recall analysis, and training/validation curves are not presented in this chapter and are therefore not treated as completed analyses here.

Overall, the experimental findings demonstrate that the hybrid YOLO–GAN model outperformed the standalone configurations across the reported evaluation metrics. The study establishes a strong basis for further development of deep learning-based automobile damage detection, with future work focusing on improving detection accuracy, evaluating performance across individual damage categories, and enhancing computational efficiency.

CHAPTER 5 - CONCLUSION AND FUTURE SCOPE

5.1 Chapter Introduction

This chapter concludes the dissertation, summarises the undertaken research, reviews and assesses the major findings and achievements of the previously established research objectives. It also presents consolidated principal contributions of the study and identifies directions for future research.

5.2 Summary of the Research

This dissertation extends the research direction established during Dissertation Phase–I, in which an integrated GAN–YOLO Model was proposed to address challenges associated with image quality, synthetic data generation, and object detection under low-resource conditions.

Dissertation Phase–II focuses on the application of this research direction to automobile damage detection. The study uses the Automobile Damage Detection dataset published on Roboflow Universe, comprising 6,957 annotated images across fourteen automobile damage categories.

The proposed Model combines the generative capabilities of GAN-based image processing with the detection capabilities of YOLO. The generative component is intended to support image enhancement and synthetic image generation, while the detection component identifies and classifies automobile damage.

The reported performance comparison considers standalone YOLO, standalone GAN, and the proposed hybrid model, with evaluation using accuracy, precision, recall, and F1-score.

5.3 Major Findings

The major research findings are mentioned below:

1. **Reported performance of the hybrid model:** The proposed YOLO–GAN model achieved a reported accuracy of 96%, precision of 93%, recall of 91%, and F1-score of 92%.
2. **Comparison with standalone Models:** The reported results show higher values across all four performance measures for the hybrid model than for the standalone YOLO and GAN Models.
3. **Standalone GAN performance:** The reported standalone GAN results are 88% accuracy, 85% precision, 82% recall, and 83% F1-score.
4. **Standalone YOLO performance:** The reported standalone YOLO results are 92% accuracy, 89% precision, 86% recall, and 87% F1-score.
5. **Application to automobile damage detection:** The research applies the hybrid Model to a dataset containing fourteen automobile damage categories, providing an application-specific context for investigating the integration of generative and detection-based deep learning.
6. **Importance of evaluation and validation:** The findings highlight the importance of consistent evaluation procedures, reliable annotations, synthetic-image validation, and further investigation of the individual contributions of enhancement and synthetic data generation.

The reported model comparison supports the conclusion that the hybrid Model achieved the strongest overall performance among the three Models presented in the source results.

5.4 Achievements during Research

The achievement of the research objectives is discussed below.

Objective 1: To design a GAN-based module for enhancing low-quality automobile images and generating realistic synthetic image samples.

Objective 2: To adapt and fine-tune a YOLO-based detector for multi-class automobile damage detection.

Objective 3: To integrate GAN-based image enhancement and synthetic image generation with YOLO-based object detection within a unified hybrid Model.

Objective 4: To experimentally evaluate the efficacy and robustness of the proposed hybrid model for automobile damage detection using appropriate performance measures.

Objective 5: To examine the computational requirements and practical applicability of the proposed Model in low-resource image-recognition environments.

5.5 Research Contributions

The primary contributions of this research are listed as follows:

1. Investigation of the hybrid GAN–YOLO model
2. Application to automobile damage recognition
3. Comparative evaluation of Models
4. Reported performance of the hybrid Model
5. Identification of practical considerations

5.6 Conclusion

This dissertation investigated a hybrid deep learning Model combining GAN-based image enhancement and synthetic image generation with YOLO-based object detection for automobile damage recognition.
The study used the Automobile Damage Detection dataset published on Roboflow Universe, comprising 6,957 annotated images across fourteen damage categories. The reported comparison included standalone YOLO, standalone GAN, and the proposed hybrid YOLO–GAN Model.

The proposed hybrid model achieved reported values of Accuracy: 96%, Precision: 93%, Recall: 91% and F1-score: 92%. The corresponding reported values for YOLO standalone were 92%, 89%, 86%, and 87%, and GAN standalone were 88%, 85%, 82%, and 83%. Thus, the hybrid Model achieved higher reported values across all four measures in the comparison.

The conclusions must be interpreted within the scope of the reported metrics and experimental setup. Overall, the research demonstrates the potential relevance of an integrated GAN–YOLO approach while identifying evaluation and optimisation required towards practical deployment.

5.7 Future Scope

The proposed research can be extended in several directions:

1. Model Optimisation
2. Real-time Deployment
3. Advanced Architectures
4. Domain-specific Applications
5. Extended Dataset Training

ACKNOWLEDGEMENT

With profound respect and heartfelt gratitude, I am sincerely grateful to the people who supported this research. Their invaluable support, input, guidance and encouragement has made this work possible.

To begin with, my deepest gratitude to Prof. Samina Mansuri, my guide, whose expertise and continuous support have been the foundation of my research work. Her insightful suggestions and patient mentorship provided the necessary direction. Her advice kept the project on track.

I am equally grateful to my co-guide, Prof. Falguni Vaghela, for her guidance and detailed feedback throughout this journey. She ensured that the quality of my work has been up to the mark that refined my core ideas of this research.

My sincere thanks to Prof. Priyanka Patel, Head of the Computer Engineering Department, for maintaining an academic environment that encouraged my growth and learning process at the institute.

I am thankful to Dr. A.C. Suthar, my esteemed principal, for his leadership at the institute and emphasis on academic excellence that has been one of my inspirations.

I would like to express my sincere gratitude to Ms. Rajlakshmi Chaudhary for her continuous encouragement, valuable suggestions, and unwavering support throughout the course of this research.

A special thanks to my family members for their unconditional love and their belief in me whenever I was facing a challenge has always been the constant support.

Finally, I appreciate everyone who offered their invaluable insights that improved this project, which is a collective effort and I remain deeply thankful to each one of you

REFERENCES

1. Hidayat, Erwin Yudi, Khafiizh Hastuti, and Azah Kamilah Muda. “Artificial intelligence in digital image processing: A bibliometric analysis.” Intelligent Systems with Applications 25 (2025): 200466.
2. Baimukashev, Daulet, Alikhan Zhilisbayev, Askat Kuzdeuov, Artemiy Oleinikov, Denis Fadeyev, Zhanat Makhataeva, and Huseyin Atakan Varol. “Deep learning based object recognition using physically-realistic synthetic depth scenes.” Machine Learning and Knowledge Extraction 1, no. 3 (2019): 883–903.
3. Asha, B., and M. Hemalatha. “Deep Learning Models for Enhancing Low-Light Photography.”
4. Farahzadi, Leila, Ibrahim Odeh, Mahdi Kioumarsi, and Behrouz Shafei. “Automated Image-Based Condition Assessment of Built Environment: A State-of-the-Art Investigation of Damage Characteristics and Detection Requirements.” Results in Engineering (2025): 104978.
5. Hassija, Vikas, Balamurugan Palanisamy, Arpita Chatterjee, Arpita Mandal, Debanshi Chakraborty, Amit Pandey, G. S. S. Chalapathi, and Dhruv Kumar. “Transformers for Vision: A Survey on Innovative Methods for Computer Vision.” IEEE Access (2025).
6. Chen, Zhaolin, Kamlesh Pawar, Mevan Ekanayake, Cameron Pain, Shenjun Zhong, and Gary F. Egan. “Deep learning for image enhancement and correction in magnetic resonance imaging — state-of-the-art and challenges.” Journal of Digital Imaging 36, no. 1 (2023): 204–230.
7. Toshpulatov, Mukhiddin, Wookey Lee, and Suan Lee. “Generative adversarial networks and their application to 3D face generation: A survey.” Image and Vision Computing 108 (2021): 104119.
8. Zhao, Feiyue, and Zhichao Zhang. “Hierarchical Graph Feature Enhancement with Adaptive Frequency Modulation for Visual Recognition.” arXiv preprint arXiv:2508.11497 (2025).

9. Zhao, Yongcan, Wei Li, Runxin Yang, and Yinghao Liu. “Real-time Efficient Image Enhancement in Low-light Condition with Novel Supervised Deep Learning Pipeline.” Digital Signal Processing (2025): 105342.
10. Umirzakova, Sabina, Shabir Ahmad, Latif U. Khan, and Taegkeun Whangbo. “Medical image super-resolution for smart healthcare applications: A comprehensive survey.” Information Fusion 103 (2024): 102075.
11. Valente, Jorge, João António, Carlos Mora, and Sandra Jardim. “Developments in image processing using deep learning and reinforcement learning.” Journal of Imaging 9, no. 10 (2023): 207.
12. Dantas, Pierre Vilar, Waldir Sabino da Silva Jr, Lucas Carvalho Cordeiro, and Celso Barbosa Carvalho. “A comprehensive review of model compression techniques in machine learning.” Applied Intelligence 54, no. 22 (2024): 11804–11844.
13. Saha, Biswanath. “The role of edge computing and cloud integration in GAN-based image synthesis.” Available at SSRN 5224776 (2025).
14. Oluwafemi, Ejiga Peter Ojonugwa, Bamidele Dayo Tunde, Mahmudul Hoque, Dapiriye Briggs, Md Mahmudur Rahman, and Fahmi Khalifa. “Colonoscopy Image Synthesis: Transforming Medical Image Classification With Enhanced Quality and Computational Efficiency.” In 2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI), pp. 1–6. IEEE, 2025.
15. Jindal, Deepanshu, Charnpreet Kaur, Ankit Panigrahi, Bani Soni, Aman Sharma, and Sanjay Singla. “Multilingual Cross-Modal Image Synthesis with Text-Guided Generative AI.” In 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT), pp. 576–582. IEEE, 2024.
16. Dunlap, Lisa, Alyssa Umino, Han Zhang, Jiezhi Yang, Joseph E. Gonzalez, and Trevor Darrell. “Diversify your vision datasets with automatic diffusion-based augmentation.” Advances in Neural Information Processing Systems 36 (2023): 79024–79034.
17. Zamir, Syed Waqas, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. “Restormer: Efficient transformer for high-resolution image restoration.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5728–5739. 2022.
18. Liang, Jingyun, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. “SwinIR: Image restoration using Swin transformer.” In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1833–1844. 2021.
19. Wang, Xintao, Yu Li, Honglun Zhang, and Ying Shan. “Towards real-world blind face restoration with generative facial prior.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9168–9178. 2021.
20. Saharia, Chitwan, Jonathan Ho, William Chan, Tim Salimans, David J. Fleet, and Mohammad Norouzi. “Image super-resolution via iterative refinement.” IEEE Transactions on Pattern Analysis and Machine Intelligence 45, no. 4 (2022): 4713–4726.
21. Wang, Xintao, Liangbin Xie, Chao Dong, and Ying Shan. “Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data.” In Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1905–1914. 2021.
22. Hussain, Muhammad. “YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection.” Machines 11, no. 7 (2023): 677.
23. Chen, Ssu-Han, and Chia-Chun Tsai. “A surface defect detection system for golden diamond pineapple based on CycleGAN and YOLOv4.” Journal of King Saud University – Computer and Information Sciences 34, no. 10 (2022): 8041–8053.

24. Jung, Do-Yoon, Yeon-Jae Oh, and Nam-Ho Kim. “A study on GAN-Based Car body part defect detection process and Comparative Analysis of YOLO v7 and YOLO v8 object detection performance.” Electronics 13, no. 13 (2024): 2598.

25. Ian, et al. “Generative Adversarial Networks.” Advances in Neural Information Processing Systems 27 (2014).

26. Automobile Damage Detection. “Automobile Damage Detection Dataset.” Roboflow Universe, October 2024. Available at: https://universe.roboflow.com/automobile-damage-detection/automobile-damage-detection/dataset/1 (accessed 2026).

IMAGE REFERENCES AND SOURCES

Figure No.	Name	References/ Sources
Fig – 2.1	Deep Learning for Image Recognition	https://canva.link/kvyluw82aqb51h8
Fig – 2.2	Image Quality Degraded vs Enhanced	https://canva.link/kvyluw82aqb51h8
Fig – 2.3	Image Enhancement Techniques	https://canva.link/kvyluw82aqb51h8
Fig – 2.4	Generative Adversarial Network Model	https://canva.link/kvyluw82aqb51h8
Fig – 2.5	YOLO based Object Detection	https://canva.link/kvyluw82aqb51h8
Fig – 3.1	Proposed Solution Flowchart	https://canva.link/de7jdr317y2sz6o
Fig – 3.2	Automobile Damage Dataset – Sample Images	https://drive.google.com/file/d/1ikVlrrfOAPh7NIdy2SIIEIPaezneSXbR/view?usp=sharing
Fig – 3.3	Data Preparation	https://canva.link/kvyluw82aqb51h8
Fig – 3.4	Data Pre-Processing	https://canva.link/kvyluw82aqb51h8
Fig – 3.5	GAN Enhanced Image	https://canva.link/kvyluw82aqb51h8
Fig – 3.6	Model Training with Hybrid Process	https://canva.link/kvyluw82aqb51h8
Fig – 3.7	Algorithmic Flowchart of Hybrid Model	https://canva.link/de7jdr317y2sz6o
Fig – 4.1	GAN Performance Chart	https://canva.link/kvyluw82aqb51h8
Fig – 4.2	GAN Image Performance Profile	https://canva.link/kvyluw82aqb51h8
Fig – 4.3	YOLO Performance Chart	https://canva.link/kvyluw82aqb51h8
Fig – 4.4	YOLO Damage Detection Profile	https://canva.link/kvyluw82aqb51h8
Fig – 4.5	Hybrid Model Image Performance Profile	https://canva.link/kvyluw82aqb51h8
Fig – 4.6	Hybrid Model Performance Chart	https://canva.link/kvyluw82aqb51h8
Fig – 4.7	Hybrid Model Process Flowchart	https://canva.link/bc4b07hmjxpbkj7
Fig – 4.8	Comparative Performance of GAN, YOLO and Proposed Hybrid Model	https://canva.link/kvyluw82aqb51h8