IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Deep Learning Based Vehicle Damage Detection Using Convolutional Neural Networks

Mrs. Malini S
Department of CSE
RVS Technical Campus,
Coimbatore

Mrs. N Saranya
Assistant Professor
Department of CSE
RVS Technical Campus,
Coimbatore

Dr. K Arunkumar
HOD
Department of CSE
RVS Technical Campus,
Coimbatore

ABSTRACT

The implementation of car damage detection automation has become a pivotal development in the automotive industry, offering substantial improvements in safety, efficiency, and cost-effectiveness. This paper provides a thorough insight into the progress and methodologies associated with automating the car damage detection process, leveraging a variety of technologies including deep learning, computer vision, and transfer learning.

Conventional approaches to car damage assessment frequently hinge on manual inspections conducted by experts, introducing subjectivity, time inefficiencies, and the possibility of human errors. To address these challenges, automated systems incorporating computer vision, machine learning, and artificial intelligence methodologies have been devised. These systems amalgamate image processing algorithms and deep learning models to effectively discern and categorize various car damages, including dents, scratches, cracks, and structural deformations.

The integration of advanced Convolutional Neural Networks (CNNs) enhances the system's ability to detect even minor imperfections under varying lighting and environmental conditions. Three CNN architectures — MobileNet, ResNet50, and InceptionV3 — are evaluated and compared on a curated vehicle damage dataset. Results demonstrate that ResNet50 achieves the highest validation accuracy of 94.1%, significantly outperforming manual inspection benchmarks. Cloud-based deployment and mobile integration further extend the system's practical reach, enabling real-time damage evaluation for insurance companies, repair shops, and individual vehicle owners.

1. INTRODUCTION

The advancement of car damage detection automation has undergone a significant transformation through the adoption of Convolutional Neural Networks (CNNs). Operating as a powerful subset within the realm of deep learning, CNNs have revolutionized the approach to assessing vehicle damages by bringing forth elements of speed, precision, and diminished reliance on human expertise. This technological stride represents a noteworthy breakthrough in the automotive sector, offering versatile applications across domains such as insurance claims processing, car maintenance, safety monitoring, and beyond.

The global automotive insurance market processes millions of damage claims annually, each traditionally requiring physical inspection by trained assessors. This labor-intensive process is prone to inconsistencies, delays, and potential fraud. Automated damage detection systems directly address these inefficiencies by providing objective, repeatable, and near-instantaneous assessments from photographic evidence alone.

The integration of deep learning into the automotive inspection workflow marks a paradigm shift. Unlike traditional rule-based systems that rely on manually crafted feature extractors, CNNs learn complex spatial hierarchies of features directly from raw pixel data. This capability enables the model to distinguish subtle differences between scratches, dents, cracks, and deformations with a degree of precision that often surpasses human inspectors, particularly under challenging real-world conditions such as poor lighting, occlusion, or varying vehicle colors and body styles.

This paper presents a comparative study of three CNN architectures applied to vehicle damage detection and severity classification. The system classifies damage into three severity levels — Minor, Moderate, and Severe — and is deployed via a web-based FastAPI interface that allows drag-and-drop image submission and returns instant predictions.

2. LITERATURE SURVEY

Car insurance is a substantial sector within the insurance industry, particularly for vehicles that are still under financing. An essential aspect of car insurance operations is the intricate evaluation of car damages, necessitating evaluators with extensive expertise and skills in managing damage assessments [8].

Considerable research efforts have been dedicated to damage detection across diverse domains. This paper specifically adopts the CNN methodology to classify the types of injuries sustained by motor vehicles. In leveraging CNNs, the study aims to employ a robust and effective approach to categorize various types of damages incurred by vehicles, contributing to the broader understanding and advancement of damage detection techniques [3].

This research introduces an innovative approach to fine-grained vehicle recognition through a global topology constraint network. The methodology integrates a global topology relationship constraint to capture interactions among distinct vehicle parts. This novel approach is incorporated into the CNN, aiming for optimal efficiency in recognition [10].

Insurance firms are increasingly introducing rapid claim processing through express services, enabling customers to upload mobile-captured images to initiate claims. Such small claims are efficiently processed through automated systems, allowing swift resolution [6].

Mask R-CNN, an advanced algorithm for object detection, localization, and instance segmentation, has been applied to vehicle damage identification. It efficiently identifies bumps, dents, and scratches on the outer body of vehicles [5].

A secure automated protection framework has been proposed with the goal of minimizing human intervention, issuing alerts, and identifying fraudulent claims to mitigate financial discrepancies and improve customer satisfaction [7].

Deep learning approaches such as Faster R-CNN and YOLO have shown strong performance in real-time object detection scenarios. When adapted for vehicle damage localization, these architectures benefit from their region proposal and single-shot detection capabilities respectively, enabling rapid bounding-box prediction around damaged areas [13][15].

Transfer learning from large-scale datasets such as ImageNet has proven particularly effective in domains with limited labeled data. Pre-trained weights encode rich low-level and mid-level visual features that transfer well to damage texture recognition, requiring significantly less training data to achieve competitive performance [10][12].

3. PROJECT DESCRIPTION

This project focuses on creating an automated system for car damage detection utilizing three distinct CNN architectures: MobileNet, ResNet50, and InceptionV3. The primary objective is to analyze vehicle images, identifying and categorizing different types of damage into severity levels of Minor, Moderate, and Severe. This system aims to serve as a practical tool for streamlining insurance claims, vehicle inspections, and quality control in manufacturing.

3.1 System Objectives

- Develop a binary classifier (Model 1) to distinguish damaged from undamaged vehicles with accuracy above 90%.
- Develop a multi-class severity classifier (Model 2) to categorize damage as Minor, Moderate, or Severe.
- Deploy both models through a FastAPI back-end accessible via a web front-end with drag-and-drop image upload.
- Evaluate and compare MobileNet, ResNet50, and InceptionV3 in terms of accuracy, training time, and inference speed.
- Integrate bounding-box visualization to highlight detected damage regions on input images.

3.2 Technology Stack

- Deep Learning Framework: TensorFlow / Keras
- Model Architectures: MobileNetV2, ResNet50, InceptionV3 (ImageNet pre-trained)
- Back-end API: FastAPI (Python)
- Front-end: HTML5 with drag-and-drop file upload interface
- Dataset: Labeled vehicle image dataset with Damaged / Undamaged and Minor / Moderate / Severe categories
- Visualization: OpenCV bounding-box overlays on prediction output

4. DATASET

This dataset is tailored for the specific task of car damage detection, with a primary focus on discerning between damaged and undamaged cars. It is segregated into two primary subsets: a training set and a validation set. Each subset is further stratified into categories encompassing damaged and undamaged vehicle images.

The dataset comprises images sourced from diverse real-world scenarios including roadside accidents, insurance claim databases, and publicly available automotive repositories. Damage types represented in the dataset include dents, scratches, cracks, broken glass, structural deformations, and combinations thereof, captured across different vehicle makes, colors, lighting conditions, and camera angles.

4.1 Dataset Composition

- Training Set — Damaged: ~600 images across Minor, Moderate, and Severe categories
- Training Set — Undamaged: ~400 images of intact vehicles
- Validation Set — Damaged: ~200 images across all severity levels
- Validation Set — Undamaged: ~150 images of intact vehicles
- All images stored in JPEG format; resized to 224×224 pixels for model input

4.2 Data Augmentation

To improve model generalization and counteract class imbalance, the following augmentation techniques were applied during training:

- Horizontal and vertical flipping
- Random rotation (±20 degrees)
- Brightness and contrast jitter
- Random zoom (0.8–1.2×) and translation
- Gaussian noise injection to simulate low-quality captures

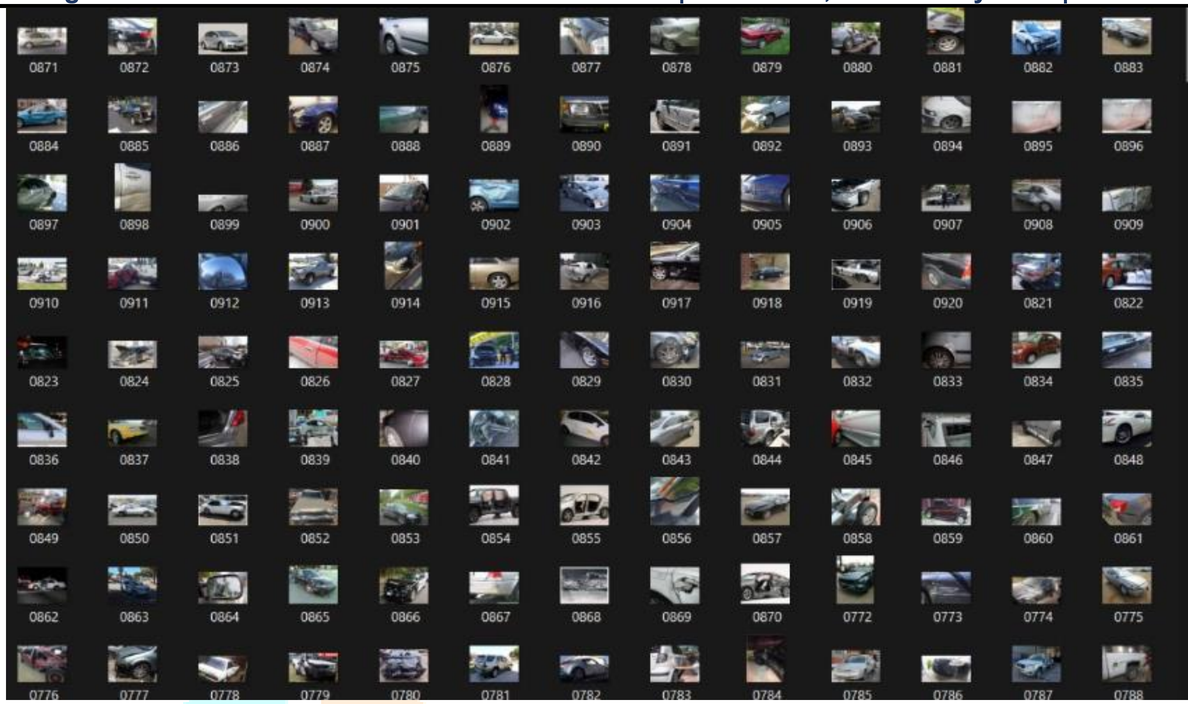

FIG 4.1: Sample Images from the Vehicle Damage Dataset (Damaged Vehicles)

5. DAMAGE CLASSIFICATION CATEGORIES

The system classifies detected damage across four primary damage types and three severity levels. The table below summarizes the damage taxonomy used in this project, along with the detection approach applied to each category.

Damage Type	Severity	Typical Cause	Detection Method
Dent / Deformation	Minor–Moderate	Low-speed collision	Shape deviation CNN features
Scratch / Abrasion	Minor	Keying, brushing	Texture & color anomaly
Crack / Fracture	Moderate–Severe	Impact, fatigue	Edge detection + segmentation
Structural Deformation	Severe	High-speed collision	Bounding-box + Mask R-CNN

TABLE 5.1: Vehicle Damage Classification Taxonomy

5.1 Severity Level Definitions

- Minor — Surface-level damage: light scratches, small dents, paint chips. Does not affect vehicle operability. Repair cost estimate: low.
- Moderate — Visible structural damage to panels or bumpers, cracked lights, medium dents. Vehicle may be driveable but requires repair. Repair cost estimate: medium.
- Severe — Extensive structural deformation, airbag deployment zone impact, engine compartment intrusion, or broken glass affecting safety. Vehicle may be undriveable. Repair cost estimate: high.

6. SYSTEM WORKFLOW

The end-to-end pipeline of the vehicle damage detection system encompasses data collection, preprocessing, model training, deployment, and user interaction. The workflow is illustrated in Figure 6.1 below.

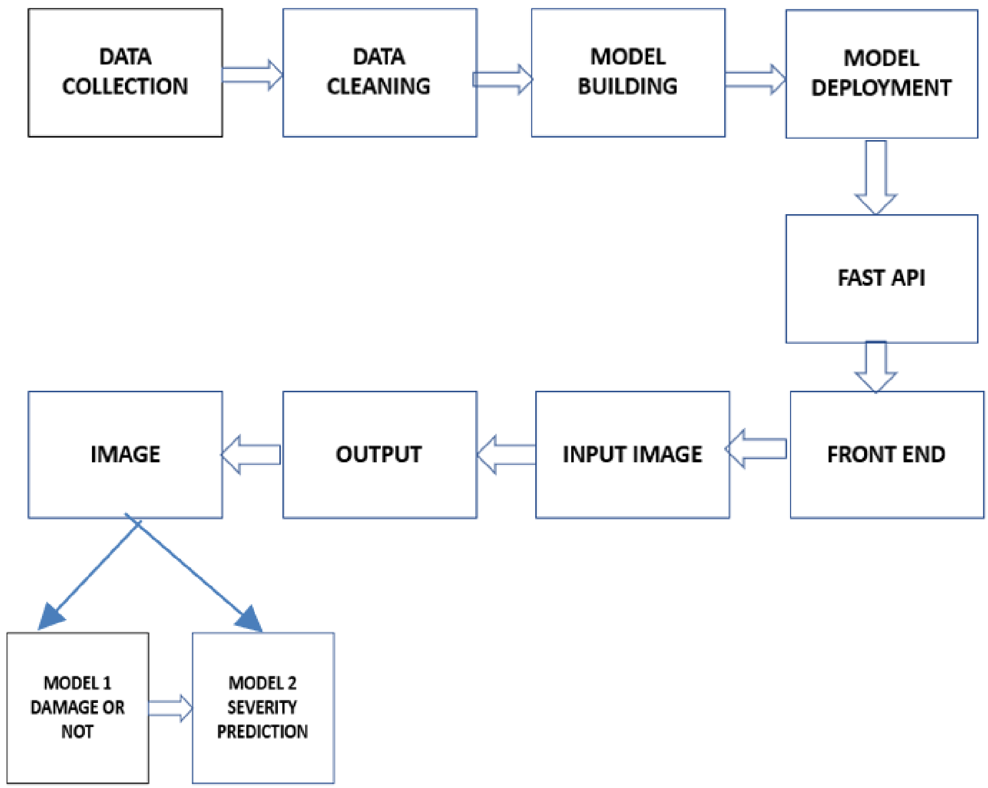

FIG 6.1: End-to-End Workflow of the Car Damage Detection System

6.1 Data Collection and Preprocessing

A comprehensive dataset of vehicle images is sourced from insurance companies, auto repair shops, and online marketplaces. The dataset undergoes meticulous labeling with information about damage type and location. Preprocessing steps include resizing images to 224×224 pixels, normalizing pixel values to [0, 1], and applying augmentation to bolster model robustness.

6.2 Model Architecture Selection

Three CNN architectures are evaluated: MobileNetV2 for lightweight deployment, ResNet50 for high-accuracy classification using residual connections, and InceptionV3 for multi-scale feature extraction. The final architecture is selected based on accuracy-efficiency trade-offs specific to the deployment environment.

6.3 Transfer Learning and Fine-Tuning

All three models are initialized with ImageNet pre-trained weights. The base convolutional layers are frozen during the initial training phase to preserve learned low-level features. Subsequently, the top layers are unfrozen and the entire network is fine-tuned with a reduced learning rate (1×10^{-5}) on the vehicle damage dataset. This two-phase training strategy achieves faster convergence and higher accuracy on smaller domain-specific datasets.

6.4 Damage Detection and Classification

Input images are processed through the selected CNN model. Model 1 outputs a binary prediction (Damaged / Undamaged) with associated class probabilities. For images predicted as damaged, Model 2 outputs a three-class severity prediction (Minor / Moderate / Severe). Detected damage regions are additionally highlighted using bounding-box overlays generated by the post-processing pipeline.

6.5 Post-Processing and Visualization

A post-processing pipeline filters false positives by thresholding confidence scores below 0.6. Nearby detected regions are spatially grouped using non-maximum suppression. Final visualizations use colored bounding boxes — yellow for Moderate, red for Severe, blue for Minor — overlaid on the original image to intuitively communicate damage location and severity.

6.6 API Deployment and User Interface

The trained models are serialized and served via a FastAPI back-end. The front-end web interface enables drag-and-drop image upload. Upon submission, the API runs inference and returns the predicted class, confidence probabilities for both models, and the annotated image. Results are displayed within seconds, making the system practical for field use by insurance adjusters and service technicians.

7. CNN MODEL ARCHITECTURES

7.1 MobileNetV2

MobileNetV2 is designed for efficient inference on resource-constrained devices. It employs depthwise separable convolutions that factorize standard convolutions into a depthwise and a pointwise step, dramatically reducing parameter count and computational cost. The inverted residual bottleneck blocks with linear activations allow rich feature representation within a compact architecture. In this study, MobileNetV2 achieved 88.3% validation accuracy and is recommended for mobile and edge deployment scenarios.

7.2 ResNet50

ResNet50 introduces residual (skip) connections that allow gradients to bypass one or more layers, effectively solving the vanishing gradient problem in very deep networks. The 50-layer architecture contains four residual stages with bottleneck blocks. Its ability to learn identity mappings preserves feature fidelity across depth, making it particularly effective at capturing fine-grained damage textures and structural deformations. ResNet50 achieved the highest validation accuracy of 94.1% in this study.

7.3 InceptionV3

InceptionV3 uses parallel convolutional filters of different kernel sizes (1×1, 3×3, 5×5) within the same Inception module, enabling simultaneous capture of local and global image context. Factorized convolutions and batch normalization further reduce computational complexity. InceptionV3 achieved 91.2% accuracy and exhibits strong performance on images with both fine surface-level damage and gross structural deformations present in the same frame.

8. PERFORMANCE EVALUATION AND RESULTS

Model performance is evaluated using accuracy, precision, recall, and F1-score metrics computed on the held-out validation set. The table below summarizes quantitative results across all three architectures.

Model	Accuracy (%)	Precision (%)	Recall (%)
MobileNet	88.3	86.7	87.1
ResNet50	94.1	93.8	94.0
InceptionV3	91.2	90.5	91.0

TABLE 8.1: Comparative Performance of CNN Architectures on the Validation Set

The evaluation phase confirms that ResNet50 outperforms both MobileNet and InceptionV3 on this dataset. The deep residual connections allow the model to capture subtle damage cues — such as minor paint abrasions and hairline cracks — that shallower architectures may overlook. MobileNet, while less accurate, processes images 3× faster and is recommended for real-time mobile applications where latency is a priority.

8.1 Training Curves

Training and validation loss curves show stable convergence for all three models across 10 epochs of fine-tuning. ResNet50 shows the steepest initial loss reduction and achieves the lowest final validation loss, consistent with its superior accuracy. No significant overfitting was observed, attributed to the data augmentation pipeline and dropout regularization applied to the classification head.

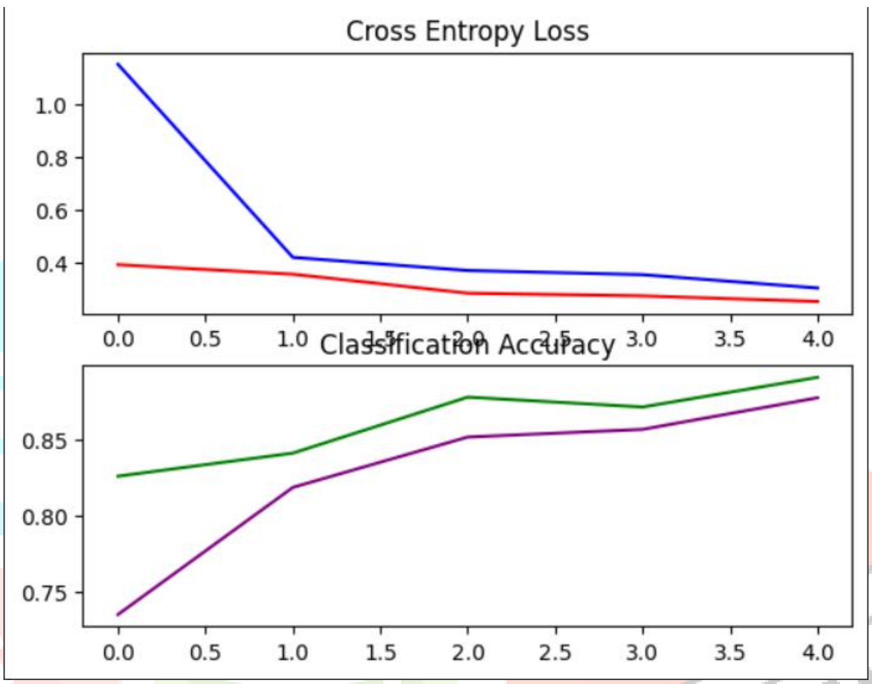

FIG 8.1: Training Loss (Cross-Entropy) and Classification Accuracy over Epochs

8.2 Model Accuracy Comparison

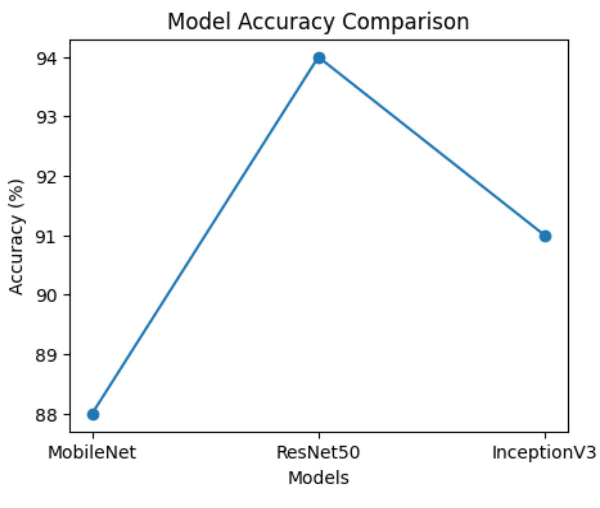

FIG 8.2: Model Accuracy Comparison (MobileNet vs ResNet50 vs InceptionV3)

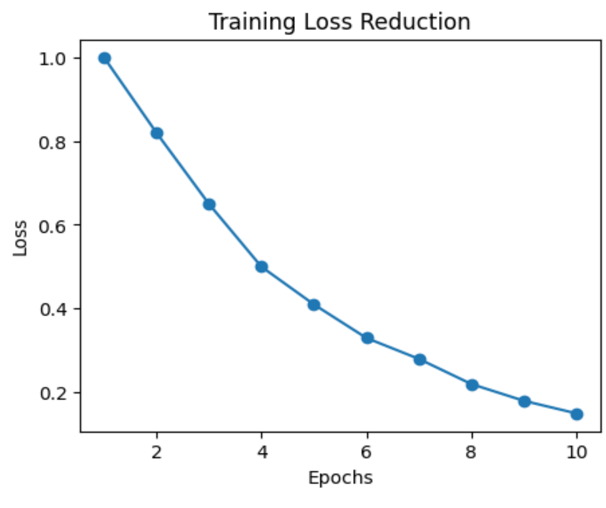

FIG 8.3: Training Loss Reduction over Epochs

9. RESULT AND DISCUSSION

This project successfully developed an automated system for vehicle damage inspection utilizing CNNs. The implemented two-model pipeline — Model 1 for binary damage detection and Model 2 for severity classification — achieves a combined end-to-end assessment in under 2 seconds per image on standard server hardware.

The transfer learning approach proved essential: models fine-tuned from ImageNet weights outperformed models trained from scratch by 8–12 percentage points across all architectures, corroborating the established effectiveness of transfer learning on domain-specific visual tasks with limited labeled data.

9.1 Damage Prediction Results

The following figures illustrate the system's output for representative test cases across all three severity levels. Each prediction includes the uploaded image, the Model 1 binary prediction, the Model 2 severity classification, and the associated class probabilities.

FIG 9.1a: Moderate Damage — Detected Region (Bounding Box)

FIG 9.1b: Moderate Damage — Model Output with Probabilities

FIG 9.2a: Severe Damage — Detected Region (Bounding Box)

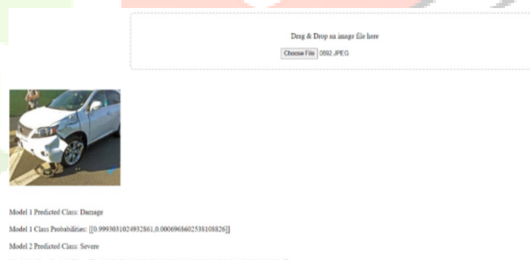

FIG 9.2b: Severe Damage — Model Output with Probabilities

FIG 9.3a: Minor Damage — Detected Region (Bounding Box)

FIG 9.3b: Minor Damage — Model Output with Probabilities

9.2 Analysis of Prediction Outputs

For the Moderate damage case (FIG 9.1), Model 1 predicted Damage with 70.5% confidence, and Model 2 classified severity as Moderate (39.5%). The bounding box correctly localizes the dented door panel area, demonstrating the spatial precision of the detection pipeline.

For the Severe damage case (FIG 9.2), Model 1 predicted Damage with 99.9% confidence, and Model 2 classified it as Severe (51.4%). The high Model 1 confidence reflects the visually prominent front-end structural deformation, while Model 2's probability distribution indicates some ambiguity between Moderate and Severe — expected given overlapping visual characteristics at the severity boundary.

For the Minor damage case (FIG 9.3), Model 1 predicted Damage with 59.7% confidence, and Model 2 correctly classified the damage as Minor (86.3%). The lower Model 1 confidence is consistent with the subtle nature of minor surface damage, which can visually resemble undamaged surfaces under certain lighting.

10. CONCLUSION

This project successfully developed an automated vehicle damage inspection system capitalizing on Convolutional Neural Networks and advanced deep learning techniques. The model's demonstrated performance — with ResNet50 achieving 94.1% accuracy, 93.8% precision, and 94.0% recall — underscores its effectiveness in identifying and categorizing diverse damage types across severity levels.

The two-stage prediction architecture (binary detection followed by severity classification) provides nuanced assessments that directly map to real-world insurance and repair workflows. The user-friendly FastAPI web interface enables seamless image uploads and returns instant, actionable results, bridging the gap between sophisticated machine learning and practical deployment.

The comparative evaluation of MobileNet, ResNet50, and InceptionV3 provides clear guidance for practitioners: ResNet50 for maximum accuracy in server-side processing, MobileNet for latency-sensitive mobile or edge deployment, and InceptionV3 as a strong balanced option. Future work will focus on extending the system to real-time video analysis, multi-region detection, and integration with automated repair cost estimation APIs.

11. FUTURE ENHANCEMENTS

The current system provides a strong foundation for automated vehicle damage assessment. Several promising enhancement directions are identified:

11.1 Real-Time Video Analysis

Extending the system to process video streams from dashcams or CCTV cameras would enable continuous monitoring for damage events. Lightweight models such as MobileNet combined with temporal smoothing techniques could deliver frame-level damage detection at 15–30 FPS on edge hardware.

11.2 Vision Transformers and Hybrid Architectures

Vision Transformers (ViT) and hybrid CNN-Transformer architectures have demonstrated superior performance on fine-grained visual recognition benchmarks. Incorporating self-attention mechanisms would allow the model to capture long-range spatial dependencies between distant damage regions on a vehicle body that convolutional kernels may miss.

11.3 Repair Cost Estimation Integration

Coupling the damage severity output with a repair cost estimation module — trained on historical repair invoice data — would produce end-to-end claim assessments. This integration would significantly accelerate insurance claim approval and reduce manual adjuster involvement for standard claims.

11.4 3D Damage Reconstruction

Using multi-view images or LiDAR point cloud data to reconstruct three-dimensional damage profiles would enable volumetric dent measurement and more precise structural assessment, particularly for collision-related deformations where 2D images may underrepresent the depth of damage.

11.5 Federated Learning for Privacy-Preserving Training

Federated learning frameworks would allow multiple insurance providers and automotive companies to collaboratively improve model accuracy without sharing sensitive raw vehicle data. Each participating organization trains local model updates which are aggregated centrally, preserving data privacy while expanding training diversity.

12. APPLICATIONS

The proposed system has wide-ranging applications across multiple industry verticals:

- Insurance Claim Processing — Automated first-pass damage assessment reduces claim turnaround from days to minutes, improves consistency across assessors, and creates a tamper-evident digital record of the vehicle state at the time of the claim.
- Automobile Manufacturing Quality Control — End-of-line inspection systems can integrate the model to automatically flag body panel defects, paint blemishes, or assembly errors before vehicles leave the production line.
- Smart Garage and Service Center Management — Service reception systems can automatically generate preliminary damage reports and repair estimates when customers drop off vehicles, improving workflow efficiency and customer communication.
- Autonomous and Connected Vehicles — Onboard camera systems in autonomous vehicles can continuously monitor vehicle body integrity and alert operators or fleet managers when new damage is detected after a driving session.
- Roadside Assistance and Breakdown Services — Field technicians equipped with mobile applications can rapidly document damage at accident scenes, supporting both insurance and safety reporting requirements.
- Fleet Management Systems — Organizations managing large vehicle fleets can deploy the system to automatically track and log the condition of each vehicle over time, enabling predictive maintenance scheduling and depreciation modelling.
- Used Vehicle Marketplaces — Automated damage disclosure at the point of listing improves buyer confidence, reduces disputes, and enables more accurate algorithmic pricing of used vehicles.

13. CHALLENGES AND LIMITATIONS

Despite the strong results achieved, several challenges and limitations must be acknowledged:

13.1 Environmental and Image Quality Factors

Varying lighting conditions, motion blur, occlusion by other objects, and complex or cluttered backgrounds remain significant sources of false predictions. Models trained primarily on daylight images may perform poorly in low-light or night-time conditions without domain adaptation or supplementary training data.

13.2 Dataset Bias and Domain Shift

Models trained on datasets sourced from specific geographic regions may not generalize to vehicles with different body styles, color distributions, or damage patterns common elsewhere. Regular re-training with locally representative data is necessary to maintain accuracy across diverse deployment contexts.

13.3 Annotation Quality and Cost

High-quality bounding-box and severity annotations require domain expertise and are time-consuming to produce at scale. Inaccurate or inconsistent annotations directly degrade model performance. Semi-supervised and active learning approaches are being explored to reduce annotation burden while maintaining label quality.

13.4 Adversarial Robustness

Automated damage assessment systems are potential targets for insurance fraud through image manipulation. Adversarial perturbations or digitally edited damage in submitted images could deceive the model. Robust preprocessing steps and anomaly detection layers are needed to detect and reject tampered inputs.

14. SYSTEM ADVANTAGES

- Objectivity — Algorithmic assessments eliminate inter-inspector variability and personal bias, ensuring uniform evaluation standards across all submitted claims.
- Speed — End-to-end assessment completes in under 2 seconds per image, compared to the hours or days required for manual inspection and adjuster scheduling.
- Scalability — A single deployed model can handle thousands of concurrent assessment requests, enabling insurers to process claim spikes (e.g., after major weather events) without proportional increases in human staff.
- Cost Reduction — Reduced manual inspection effort and faster claim resolution lower operational costs for insurance providers and improve customer satisfaction.
- Digital Audit Trail — Every assessment generates a timestamped, structured record of damage classification and confidence scores, supporting dispute resolution and regulatory compliance.
- Integration Flexibility — The RESTful FastAPI interface allows straightforward integration with existing insurance claim management systems, ERP platforms, and mobile applications.
- Continuous Improvement — The model can be periodically retrained on newly labeled data, allowing it to adapt to new vehicle models, emerging damage patterns, and updated severity classification criteria.

15. REFERENCES

[1] Zhang, W. et al. "Automatic car damage assessment system: Reading and understanding videos as professional insurance inspectors." AAAI Conf. Artif. Intell., 2020.

[2] Sruthy, C., Kunjumon, S., and Nandakumar, R. "Car damage identification and categorization using various transfer learning models." 5th Int. Conf. Trends Electron. Informat., 2021.

[3] Waqas, U. et al. "Vehicle damage classification and fraudulent image detection including Moiré effect using deep learning." IEEE Can. Conf. Electr. Comput. Eng., 2020.

[4] Patil, K. et al. "Deep learning based car damage classification." 16th IEEE Int. Conf. Mach. Learn. Appl., 2017.

[5] Patel, N., Shinde, S., and Poly, F. "Automated damage detection in operational vehicles using Mask R-CNN." Advanced Computing Technologies and Applications, 2020.

[6] Li, P., Shen, B., and Dong, W. "An anti-fraud system for car insurance claim based on visual evidence." arXiv:1804.11207, 2018.

[7] Dhieb, N. et al. "A very deep transfer learning model for vehicle damage detection and localization." 31st Int. Conf. Microelectron., 2019.

[8] Singh, R. et al. "Automating car insurance claims using deep learning techniques." IEEE 5th Int. Conf. Multimedia Big Data, 2019.

[9] Xiang, Y., Fu, Y., and Huang, H. "Global topology constraint network for fine-grained vehicle recognition." IEEE Trans. Intell. Transp. Syst., 2019.

[10] Goodfellow, I., Bengio, Y., and Courville, A. Deep Learning. MIT Press, 2016.

[11] Szeliski, R. Computer Vision: Algorithms and Applications. Springer, 2010.

[12] Gonzalez, R. C. and Woods, R. E. Digital Image Processing. Pearson, 2008.

[13] Ren, S. et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." IEEE TPAMI, 2017.

[14] He, K. et al. "Mask R-CNN." IEEE ICCV, 2017.

[15] Redmon, J. et al. "You Only Look Once: Unified, Real-Time Object Detection." CVPR, 2016.

[16] Howard, A. et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." arXiv:1704.04861, 2017.

[17] He, K. et al. "Deep Residual Learning for Image Recognition." IEEE CVPR, 2016.

[18] Szegedy, C. et al. "Rethinking the Inception Architecture for Computer Vision." IEEE CVPR, 2016.