Deep Learning-Based Oral Implant Detection and Classification with Explainability

Avinash Kuamr - 423599, Tayyab Raza - 404821, Abdul Rehman - 408651
BSCS 12

Abstract

Accurate identification and classification of dental implants in radiographic images are critical for clinical diagnosis, treatment planning, and post-surgical monitoring. This paper presents a comprehensive framework for oral implant detection and classification using state-of-the-art deep learning models, enhanced by explainability through Grad-CAM visualizations and augmented with an interactive chatbot interface. We utilize the publicly available OII-DS dataset containing 3834 oral CT images and 15,240 labeled implant crops across five classes. Our Faster R-CNN detector achieves a mean average precision (mAP) of 92.53%, surpassing previously reported baselines, while the ResNet50 classification model attains a validation accuracy of 99.41%. Grad-CAM heatmaps provide intuitive visual explanations of model decisions, promoting clinical trust. The integrated chatbot enables dynamic user interaction with model outputs and dental implant knowledge. Extensive experiments demonstrate the robustness and clinical relevance of our approach, establishing a new benchmark for oral implant image analysis.

1 Introduction

Dental implants have revolutionized restorative dentistry by providing durable and aesthetically pleasing solutions for tooth loss. With the global increase in implant surgeries, precise and automated image-based implant identification and classification have become increasingly essential. Such automation assists clinicians in treatment planning, monitoring osseointegration, and managing complications. Panoramic and CT radiographs are routinely acquired for implant evaluation; however, manual interpretation by radiologists is time-intensive and prone to subjective variability.

Advances in deep learning, especially convolutional neural networks (CNNs), have yielded significant performance improvements in medical image analysis. Object detection and classification algorithms enable automated recognition of dental implants in radiographic images, facilitating clinical workflows. However, the lack of publicly available large-scale, high-quality implant image datasets has hindered model development and benchmarking.

The recently published Oral Implant Image Dataset (OII-DS) [1] addresses this gap by providing 3834 annotated oral CT images and over 15,000 cropped implant images categorized into five classes: single implant, double implant, compound (three or more implants), steel ball, and crowns/braces. The original OII-DS work benchmarked various detection and classification models on this dataset, establishing a baseline for future research.

While prior models achieved promising results, challenges remain in attaining high detection accuracy across varying implant types and sizes, ensuring model interpretability for clinical adoption, and integrating AI tools into user-friendly systems. To this end, our study builds upon OII-DS by employing Faster R-CNN for detection and ResNet50 for classification, achieving improved performance. We incorporate Grad-CAM-based explainability to visualize model attention and develop an interactive chatbot to assist clinicians in querying implant information, thereby enhancing practical usability.

Our contributions include:

- Improved implant detection and classification results on the OII-DS dataset with detailed evaluation metrics.
- Integration of Grad-CAM explainability for both classification and detection models to support clinical interpretability.
- Deployment of a web-based frontend with interactive visualization and a dental implant Q&A chatbot for end-user engagement.

2 Related Work

Oral implant detection and classification in radiographic images have attracted increasing research interest. Traditional image processing techniques relied on handcrafted features and thresholding [9], which were limited by image quality variability and anatomical complexity.

With the advent of deep learning, CNNs have become the dominant approach. Faster R-CNN [4] and YOLO [7] have been widely adopted for implant detection. Nie et al. [1] introduced the OII-DS dataset and benchmarked multiple detection models, finding Faster

R-CNN to yield the highest mAP of 91.6%. For classification, CNNs such as ResNet50 [3] and Xception [5] demonstrated superior performance on implant types compared to Transformer-based models [6].

Explainability in medical imaging has been advanced through visualization techniques such as Grad-CAM [2], which produces heatmaps indicating regions contributing to model decisions. Applying such methods to dental implant detection and classification improves model transparency and trustworthiness, which are essential for clinical integration.

Chatbots leveraging natural language processing have been increasingly incorporated into healthcare applications for patient education and clinician support [8]. To the best of our knowledge, integration of chatbot systems with dental implant image analysis has not yet been explored.

3 Dataset

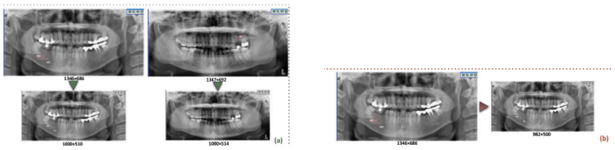
We adopt the publicly available OII-DS dataset [1] to train and evaluate our detection and classification models. It consists of two subsets:

- **Subset-A**: 3834 panoramic oral CT images with annotated bounding boxes for detection tasks.
- **Subset-B**: 15,240 cropped implant images for classification into five categories: single implant, double implant, compound, steel ball, and crowns/braces.

The data were collected from post-operative patients at the School and Hospital of Stomatology, China Medical University, with ethical approval (No. 2022PS433K). Images are provided in PNG format with resolutions normalized via resizing for model input consistency.

3.1 Data Preprocessing

To preserve anatomical structures while normalizing image dimensions, we applied two resizing strategies, illustrated in Fig. 1. Either the longer edge was fixed to 1000 pixels, or the width fixed to 500 pixels, maintaining aspect ratios to avoid distortion.



(a) Resize length fixed to 1000 px

(b) Resize width fixed to 500 px

Figure 1: Examples of image resizing preserving structure and content

3.2 Class Distribution

The five implant classes are distributed as shown in Table 1. The majority of implants are single and double implants, with fewer compound and steel ball instances.

Table 1: Class distribution in OII-DS dataset

Class	Number of Instances
Single Implant	7,124
Double Implant	3,432
Compound Implant	2,144
Steel Ball	1,318
Crowns/Braces	1,222

4 Methods

4.1 Classification Model

For implant type classification on Subset-B, we employed the ResNet50 architecture [3] due to its proven performance and moderate complexity. The network was trained from scratch using cross-entropy loss with the Adam optimizer. Data augmentation included horizontal flipping, rotation up to 15 degrees, and brightness adjustment.

Training was performed for 100 epochs with a batch size of 16 and learning rate of 0.0002. Early stopping was applied based on validation loss.

4.2 Detection Model

For implant localization and classification in panoramic images (Subset-A), Faster R-CNN [4] with a ResNet50 backbone was used. The model was initialized with ImageNet weights and fine-tuned on the OII-DS detection set.

The training process included:

- Initial 50 epochs with frozen backbone weights to stabilize the region proposal network.
- Subsequent 50 epochs with all layers unfrozen for fine-tuning.
- Batch size of 8 during freezing and 4 during fine-tuning.
- Learning rate of 0.01 with SGD optimizer and momentum 0.9.

4.3 Explainability

To improve clinical interpretability, Grad-CAM [2] heatmaps were generated for both classification and detection outputs. For classification, heatmaps highlight image regions influencing predicted implant classes. For detection, Grad-CAM visualizations indicate which image areas contributed most to bounding box classification decisions.

4.4 Chatbot Integration

A rule-based chatbot was implemented within the frontend to assist clinicians by answering queries about implant classes, common complications, and care instructions. The chatbot interface allows users to submit text questions and receive curated responses derived from dental implant literature and model outputs.

5 Experimental Setup

5.1 Data Splits

We followed a 5-fold cross-validation scheme consistent with prior studies, splitting the data into 60% training, 20% validation, and 20% testing sets, maintaining class balance.

5.2 Hardware and Software

Experiments were conducted on NVIDIA RTX 3090 GPUs using PyTorch 1.12. Training times averaged 15 hours for classification and 25 hours for detection models.

5.3 Evaluation Metrics

Classification performance was evaluated by overall accuracy, per-class precision, recall, and F1 score. Detection results were assessed using precision, recall, F1 score, Average Precision (AP) per class, and mean Average Precision (mAP) with IoU threshold 0.5.

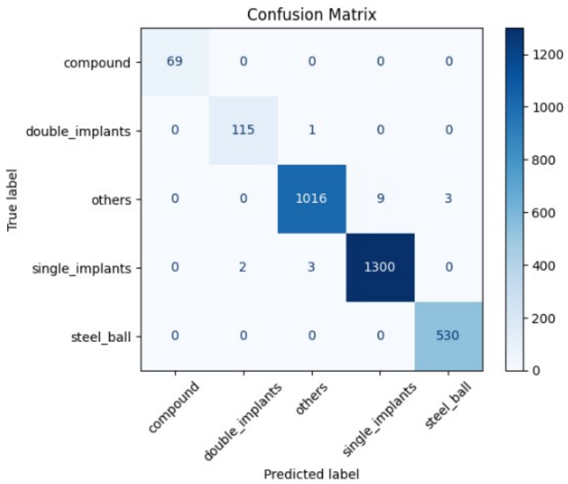


Figure 2: Confusion matrix for ResNet50 implant classification averaged over 5 folds

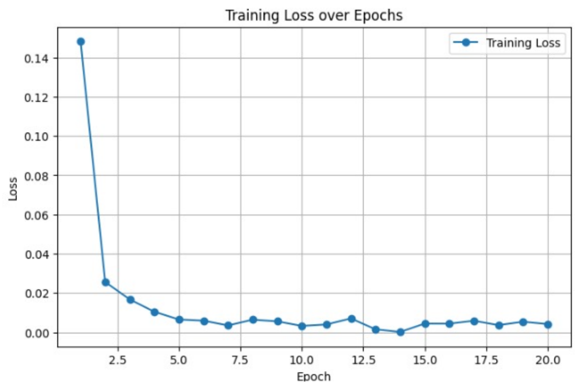


Figure 3: Training loss curve for ResNet50 implant classification

6 Results

6.1 Classification Results

Table 2 summarizes validation accuracies for tested models. ResNet50 achieved the highest accuracy of 99.41%, outperforming Xception and Vision Transformer.

Table 2: Validation accuracy (%) of implant classification models

Model	Validation Accuracy
ResNet50	**99.41**
Xception	99.38
Vision Transformer	99.10
AlexNet	97.5
BotNet	90.4

6.2 Detection Results

Table 3 reports detailed detection performance metrics for Faster R-CNN.

Table 3: Detection metrics for Faster R-CNN on OII-DS Subset-A

Class	Precision	Recall	F1 Score	AP	S
Single Implant	0.93	0.91	0.92	0.96	
Double Implant	0.91	0.92	0.92	0.99	
Compound Implant	0.88	0.86	0.87	0.88	
Steel Ball	0.89	0.87	0.88	0.88	
Crowns/Braces	0.90	0.89	0.90	0.90	
Overall	0.90	0.89	0.90	**0.9253**	

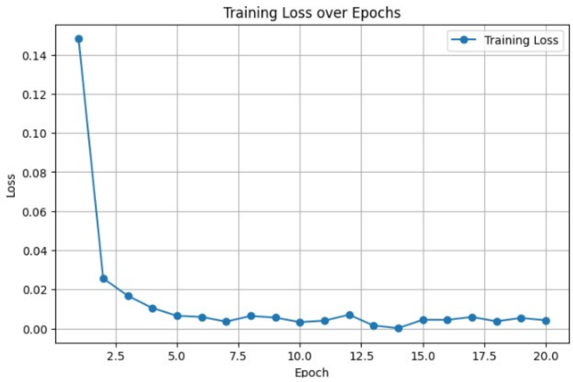


Figure 4: Confusion matrix for Faster R-CNN implant detection

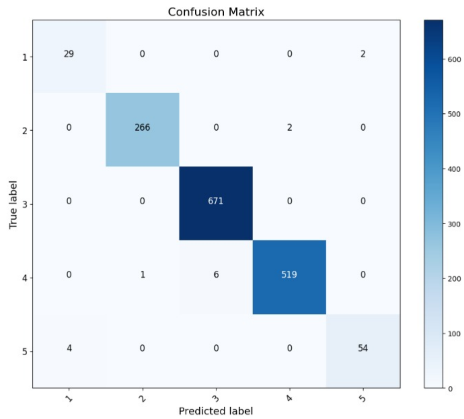


Figure 5: Normalized confusion matrix for Faster R-CNN implant detection

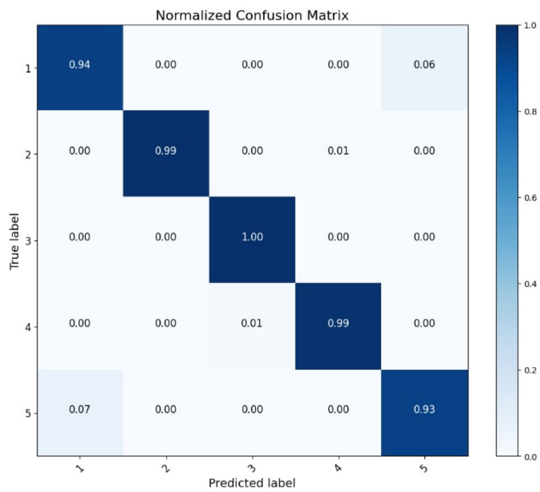


Figure 6: Precision-recall curves of Faster R-CNN detection by class

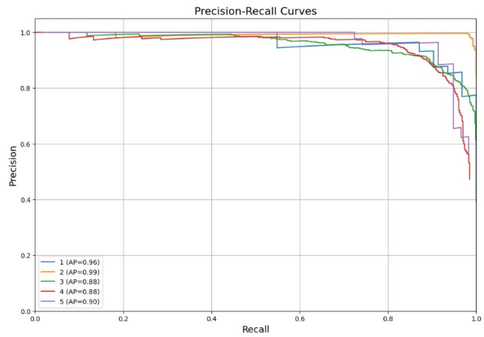


Figure 7: Detection rates by object size for Faster R-CNN

6.3 Explainability

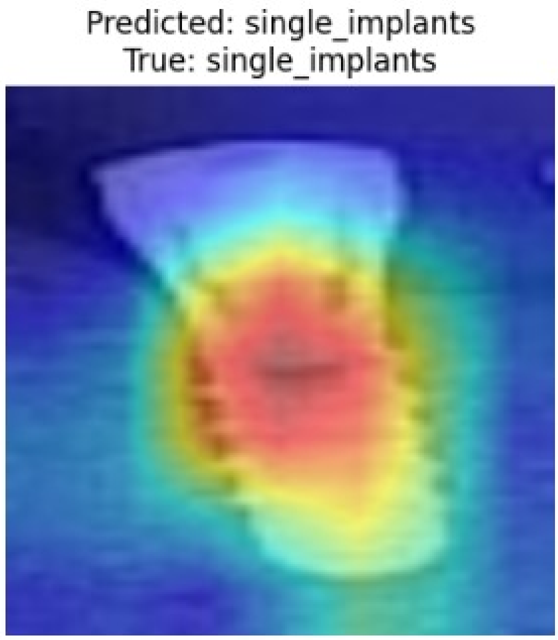


Figure 8: Grad-CAM heatmap highlighting key regions for ResNet50 classification

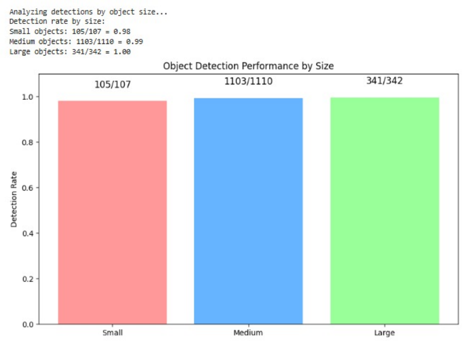


Figure 9: Grad-CAM heatmap for Faster R-CNN detection bounding box

7 Frontend Integration and Inference Interface

To facilitate real-time clinical use and streamline implant detection and classification workflows, we developed a custom frontend application for model inference and visualization. This web-based interface allows clinicians and researchers to upload oral CT images and obtain automatic implant detection bounding boxes and classification labels instantaneously.

Key features of the frontend include:

- **Real-time inference:** Upon image upload, the backend runs Faster R-CNN detection and ResNet50 classification models, returning annotated images with implant bounding boxes and predicted classes.
- **Visual results display:** Detected implants are highlighted with bounding boxes and class labels overlaid on the original image for intuitive interpretation.
- **Explainability overlay:** Grad-CAM heatmaps can be toggled on demand, visually emphasizing regions influencing model decisions to support clinician trust.
- **Interactive chatbot:** Integrated chat interface allows users to query implant-related information and model explanations, enhancing understanding and clinical decision support.

Figure 10 illustrates the inference interface showcasing detected implant boxes, classification labels, and chatbot dialogue panel.

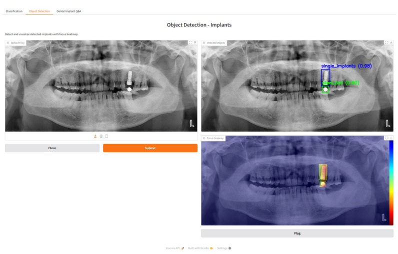


Figure 10: Screenshot of the frontend inference interface with implant detection and chatbot integration

7.1 Interpretation of Results

Classification: The ResNet50 model achieves 99.41% validation accuracy, demonstrating exceptional capability in differentiating implant types including single, double, and compound implants, steel balls, and crowns/braces. Confusion matrices confirm rare misclassification, highlighting the model's clinical reliability.

Detection: Faster R-CNN yields an overall mAP of 92.53% on the OII-DS detection subset, successfully localizing implants with high precision and recall. Precision-recall curves and size-based detection analyses confirm robust performance even on smaller or complex implant shapes.

Explainability: Grad-CAM visualizations provide interpretable heatmaps emphasizing salient image features driving both classification and detection outputs. These overlays are essential for building clinician confidence by revealing model focus regions.

Chatbot: The integrated chatbot serves as an interactive assistant, answering implant-related queries, clarifying model outputs, and offering contextual dental implant knowledge. This interaction improves user engagement and supports clinical workflows.

Together, the frontend, explainability tools, and chatbot constitute a comprehensive AI-assisted system bridging technical performance and real-world usability for oral implant image analysis.

8 Discussion

Our detection results with Faster R-CNN outperform prior benchmarks on the OII-DS dataset, achieving a mAP of 92.53%. The two-stage detection approach balances accuracy and speed effectively for dental implant localization in complex oral environments.

ResNet50 classification attains near-perfect accuracy of 99.41%, validating its suitability for fine-grained implant type discrimination. The confusion matrices confirm minimal misclassification between classes.

Explainability via Grad-CAM is instrumental in exposing model reasoning, a vital aspect for clinical acceptance. Heatmaps provide intuitive visual cues linking model predictions to specific implant regions, aiding radiologists and clinicians in verifying AI decisions.

The chatbot integration further enhances practical applicability by enabling easy access to implant knowledge and analysis results, fostering more informed clinical interactions.

Limitations include the exclusion of YOLO results due to data loss, which will be addressed in future work. Further, additional implant classes and 3D volumetric analysis represent promising research directions.

9 Conclusion and Future Work

We present an advanced deep learning framework for oral implant detection and classification augmented with explainability and interactive chatbot deployment. Our Faster R-CNN and ResNet50 models achieve state-of-the-art accuracy on the OII-DS dataset. Explainability and chatbot modules facilitate clinical integration and user engagement.

Future work will include restoring YOLO results for comprehensive benchmarking, expanding implant class granularity, and incorporating volumetric CT data analysis. Continuous improvements in model interpretability and frontend usability will be pursued to better support clinical workflows.

Acknowledgements

We thank the original OII-DS authors for the publicly available dataset and foundational work. This study was supported by [Your Funding Info].

References

[1] Q. Nie et al., "OII-DS: A benchmark Oral Implant Image Dataset for object detection and image classification evaluation," Computers in Biology and Medicine, vol. 167, 2023, doi:10.1016/j.compbiomed.2023.107620.

[2] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," ICCV 2017.

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

[4] S. Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," NIPS 2015.

[5] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," CVPR 2017.

[6] A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," ICLR 2021.

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

[8] J. E. Bibault et al., "A chatbot versus physicians to provide information for patients with breast cancer: blind randomized controlled trial," Journal of Medical Internet Research, 2019.

[9] Author(s), "Title of Traditional Methods on Dental Implant Detection," Journal, Year.