

Received 27 March 2025, accepted 10 April 2025, date of publication 15 April 2025, date of current version 22 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3560998



RESEARCH ARTICLE

Fusion of Image Filtering and Knowledge-Distilled YOLO Models for Root Canal Failure Diagnosis

AFIFA ZAIN APURBA^{ID}, MD FAHIM SHAHORIAR TITU^{ID}, MAHIR AFSER PAVEL^{ID}, INTISAR TAHMID NAHEEN^{ID}, AND RIASAT KHAN^{ID}, (Member, IEEE)

Department of Electrical and Computer Engineering, North South University, Dhaka 1229, Bangladesh

Corresponding author: Riasat Khan (riasat.khan@northsouth.edu)

ABSTRACT Root canal treatment involves the removal of inflamed pulp from infected teeth and sealing them to prevent re-entry of bacteria and infection. Early and accurate identification of root canal issues is crucial for improving treatment outcomes and minimizing complications. This study proposes a novel method for efficiently detecting failures in root canal treatments by incorporating image filtering techniques into deep learning models. A custom dataset has been collected from Square Hospital Ltd of Dhaka, Bangladesh, comprising 1,600 annotated, filtered radiographs. These images are processed using denoising algorithms (mean, median, Gaussian, contourlet transform, or Bayesian wavelet), non-local means, and BM3D. The YOLO models used for detection included YOLOv5, YOLOv7, and YOLOv8, with YOLOv5 achieving the highest accuracy of 90.6% on the filtered datasets. This work integrates an autodistillation technique with Grounded SAM as the base model for bounding box generation, YOLOv8 as an intermediate model to bridge architectural differences, and YOLOv5 as the target model due to its lightweight design, fewer parameters, and superior performance on the filtered dataset. Knowledge distillation was employed to optimize performance on lightweight models, resulting in precision and recall values of 99.16% and 97.29%, respectively, for YOLOv5. The filtered datasets outperformed raw images, with the Bayesian wavelet and contourlet transform yielding the best results on YOLOv5. This pipeline demonstrates the potential to develop one of the most accurate and clinically applicable diagnostic tools for root canal failure detection by combining deep learning and knowledge distillation techniques with image filtering. This study demonstrates how combining deep learning with advanced image enhancement techniques can provide a highly accurate, clinically relevant tool for the early detection of root canal treatment failures, potentially aiding dentists in making faster and more consistent diagnoses.

INDEX TERMS Root canal failure detection, filtered radiographs, YOLO models, knowledge distillation, image filtering techniques, deep learning in dentistry.

I. INTRODUCTION

Oral diseases are becoming increasingly common due to modern diets and other factors, with endodontic issues being a significant concern. According to the American Association of Endodontists, approximately 25 million root canal procedures are performed annually, averaging about 41,000 per day or 287,000 per week [1]. Root canal treatment

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy^{ID}.

involves removing the inflamed or infected pulp, cleaning and disinfecting the tooth, and sealing it with a cap to eliminate bacteria and prevent reinfection. This treatment is necessary when a cavity is too large and has reached the soft tissues of the tooth, causing pain when chewing or biting. An infected tooth can often be saved if diagnosed early enough to prevent more complex issues later [2]. Symptoms of infection include pain, swelling, loose gums, tooth discoloration, and loosening of the tooth due to the softening of the surrounding bone [3].

Traditionally, dentists diagnose dental problems by evaluating X-ray films. However, this approach is prone to errors due to technical issues, human errors and variations in light exposure, which can result in inaccurate human assessments. The need for dentists to perform lengthy procedures makes it challenging to identify healthy teeth suspected of infection accurately. Deep learning technology has significantly reduced errors at various stages, saving both time for patients and physicians. When performed accurately, AI-based X-ray analysis of infected teeth can determine whether an infected tooth can be saved or needs to be extracted. This technique leads to more cost-effective and faster treatments for patients [4].

Image filtering is one of the fundamental techniques in image processing, aimed at improving the quality of X-ray imagery by preserving all relevant information [5]. This technique eliminates noise, enhances feature detection, and normalizes variations in radiographic images, resulting in consistent outcomes within each filter category [6]. The output of the filter depends on parameters such as the shape and size of the filter kernel. Research in AI applications for oral diseases has shown promising results. For example, Kaya et al. focused on detecting and classifying dental structures in pediatric radiographs using YOLOv4 and CNN-based model designs [7]. Li et al. developed a tooth position detection method using Gaussian filtering and YOLOv4 techniques [8], while Chen et al. proposed registration methods for 3D detection of tooth cracks [9].

Interestingly, these studies have provided some of the most relevant approaches to diagnosing and detecting features in oral health care. However, they did not apply filtering methods to improve performance, did not incorporate current or state-of-the-art techniques, and lacked sufficient normalization for model comparisons. Presented in Figure 1 is a mind map that illustrates the core concepts, relationships, and ideas explored in this research, serving as a visual representation of the study's framework and key areas.



FIGURE 1. Core areas of focus in dental imaging solutions.

An advanced AI application in dentistry employs neural network models to identify various stages of root canal treatment. The proposed model compares the performance of raw and denoised images and aims to develop a reliable diagnostic tool that overcomes the limitations of traditional methods. This work promotes early detection and effective intervention by designing a neural network model specifically to identify the stages of root canal treatment.

This work aims to develop a reliable diagnostic tool to improve early detection and outcomes of interventions in dentistry, demonstrating the following contributions.

- A significant contribution of this work is to present a private dataset comprising 1,600 radiographic images collected from a local hospital in Dhaka, Bangladesh, representing various stages of root canal treatment.
- Labeling of the images is performed using tools such as LabelImg for three-class images, with problematic teeth annotated using bounding boxes.
- Denoising techniques, e.g., Mean, Median, Gaussian, and BM3D are applied to enhance image clarity and facilitate annotation. During training, both filtered and unfiltered images are compared.
- Deep learning models (YOLOv5, YOLOv7, YOLOv8) are evaluated for reliability across metrics, including precision, recall, mAP, and accuracy.
- A comparison of filtered images is conducted across models, and knowledge distillation employing Ground SAM, YOLOv8 and YOLOv5 models is applied to optimize performance. In the context of hierarchical knowledge distillation, we aim to transfer the learning from complex, large models to more lightweight models as efficiently as possible.

The novelty of this research lies in the application of various filtering techniques and knowledge distillation-based deep learning methods to a private dataset consisting of 1,600 radiographic root canal images.

II. RELATED WORKS

Radiographs are widely used for medical purposes, leading to an increase in the number of articles focused on X-ray detection. With advancements in artificial intelligence, computer vision and deep learning techniques have been applied in numerous studies for object detection and image segmentation. We describe some recent articles on radiographic detection and root canal identification using deep learning techniques.

Table 1 summarizes the key methodological challenges, computational constraints, and interpretability issues of recent studies on automatic radiographic detection. The literature reviews highlight the extensive research on identifying various dental issues using deep learning techniques. However, several notable drawbacks remain in the use of advanced deep learning methods for detecting and classifying dental conditions. Earlier studies were primarily limited by smaller datasets, the absence of applying lightweight models, fewer experimental methods, and the absence of filtering or advanced techniques to improve results, as addressed in this study.

III. MATERIALS AND METHODS

A. RESEARCH DESIGN

A comprehensive system has been implemented in this study utilizing deep learning models to effectively classify and detect root canal treatments. The process involves several steps, from data collection to model evaluation. The overall

TABLE 1. Summary of studies on radiographic detection using deep learning techniques.

Study	Research Focus	Dataset	Methodology	Major Contribution	Main Findings	Limitations
Radiographic Detection						
[10]	Wrist injury detection	435 radiographic images	Two-stage neural networks (Object detection, Binary classification)	A deep learning model for detecting perilunate and lunate dislocations introduced.	High performance: Accuracy (93.4%), F1-score (93.8%), AUC (0.986)	The subtle or overlapping injuries remain to be verified for detection by the model, along with lack of interpretability in a clinical diagnosis.
[11]	Chest disease detection	26,316 chest X-ray images	Deep learning (DC-ChestNet, VT-ChestNet)	Proposed two-step deep learning for ETDD classifying chest disease.	Achieved AUC: 99.26% (DC-ChestNet) and 99.57% (VT-ChestNet)	Though two-staged processing increases computational complexity, the patient's limited model transparency affects clinical trust.
[12]	COVID-19 detection	Chest X-rays	DarkNet model	DarkNet: classified COVID-19 cases apart from healthy and pneumonia individuals.	Accuracy: 87.06% (multiclass), 97.88% (binary)	Multiclass classification tasks already have a big backlash to the model performance, which does not include region-specific details in clinical validation.
[13]	Automated chest disease detection	112,120 chest X-rays	CNN models	Created an automated pipeline for multiple chest disease diagnosis.	High AUC for hernia (0.945) and emphysema (0.912)	The pipeline is limited in scope with regard to imaging devices. Great computational resources are needed to run it.
[14]	Defect inspection	Not Specified	Novel deep learning models	Presented an exhaustive review on the advancements and challenges of deep learning in defect inspection.	Summarized future directions and applications	The review material does not provide experimental validation nor does it treat challenges related to real defect detection.
[15]	Automatic diagnosis of COVID-19 through radiographic imaging.	Public chest X-ray and CT datasets	Explainable Attention Transfer Classification Model Using a Knowledge Distillation Network Architecture.	A new explainable paradigm is proposed for detecting COVID-19, which involves the use of attention transfer and deformable attention modules in a model architecture and also involves an image fusion module that is engineered to increase the performance of the model further.	The model was demonstrated to differentiate COVID-19 infection from community-acquired pneumonia and healthy lungs while improving the interpretability as well.	Evaluation was limited across demographics and imaging conditions, and it was not fully validated for real-world clinical application.
Tooth Detection						
[16]	Root dilacerations detection	636 panoramic radiographs (983 objects)	Deep learning models	Evaluated deep learning's efficacy in root dilaceration diagnosis	mAP: 0.92, Accuracy: 0.72, F1-score: 0.83	The quality of radiograph hen detection of artifacts becomes susceptible will play a big role in the accuracy.
[17]	Teeth detection and numbering	1352 panoramic radiographs	CNN (VGG16)	CNN implemented for automatic dental numbering.	F1-score: 0.998, Precision: 0.9998	The pipe architecture of VGG16 is not flexible enough to accommodate complex dental anomalies and as yet has not ruled out issues concerning model interpretability.
[18]	Root canal complication detection	915 dental radiographs	CNN, LSTM	The main contribution was the comparison of the architectures using CNN and LSTM for detecting root canal issues.	Filtered CNN model performed best	This study is not comprehensive with respect to evaluation metrics, and the comparison of CNN and LSTM needs to be validated clinically.
[7]	Pediatric dental structure detection	4821 pediatric radiographs	YOLOv4, CNN	It has developed models for the detection of immature dental structures.	Promising diagnosis of immature teeth	The computational overhead of YOLOv4 makes it not clinically applicable and the performance inconsistency needs to be explored across age groups.
[8]	Tooth position identification	Not specified	Gaussian filtering + YOLOv4	Combined Gaussian filtering with YOLOv4 for segmentation of teeth parts.	Achieved 90% accuracy	Gaussian filtering might introduce noise and the potential clinical applicability of YOLOv4 in real-world dental cases is less.
[9]	3D tooth crack detection	Not specified	3D deep learning models	Designed a 3D visualization model to identify tooth cracks.	98% similarity with microscopic inspection	High computation and complexity of the model render it unfeasible for real-time applications.
[19]	Deep Learning Automated Endodontic Difficulty	Dataset: 1,386 annotated periapical radiographs acquired from two clinical sites.	CNNs (e.g., VGG16, ResNet) with transfer and self-supervised learning, evaluated via 10-fold cross-validation.	Validation of deep learning for automated difficulty evaluation.	VGG16 reached an accuracy of 87.62%; models outshone human raters; self-supervised learning did not show improvement.	Self-supervised pretraining did not improve performance; the limited small size of the dataset and possibly biased annotations.

workflow is illustrated in Figure 2, which summarizes various activities undertaken in this study.

An extensive custom dataset was compiled by collecting radiographic images of various stages of root canal treatments. All images were sorted and selected based on clear visuals to meet the requirements of deep learning models and to standardize the data. Contouring was performed to mark the affected tooth, while extraneous parts were cropped and resized. Null or defective images were identified and removed. Additionally, normalization techniques were applied to enhance the intensity range of the images, ensuring that the dataset predominantly consisted of high-quality data for research.

The annotated dataset was created in two phases. In the first phase, selected images were manually annotated by

professional dentists. In the second phase, the same images were auto-annotated using a base model from the autodistill process. This model identified and classified the affected tooth using bounding boxes. After annotation, several denoising algorithms, including Gaussian, Bayesian wavelet, mean, and median filtering were applied to the images to improve clarity. Denoising filters were applied to make X-ray images clearer, helping the AI model detect treatment issues more accurately, i.e., just as a clearer image would help a dentist spot problems faster. Some noisy images were left untouched for comparison.

All datasets underwent various augmentation procedures, including cropping, resizing, rotation, and flipping, to expand and balance the dataset. The final dataset was divided into a learning dataset (80%) and a testing dataset (20%). The

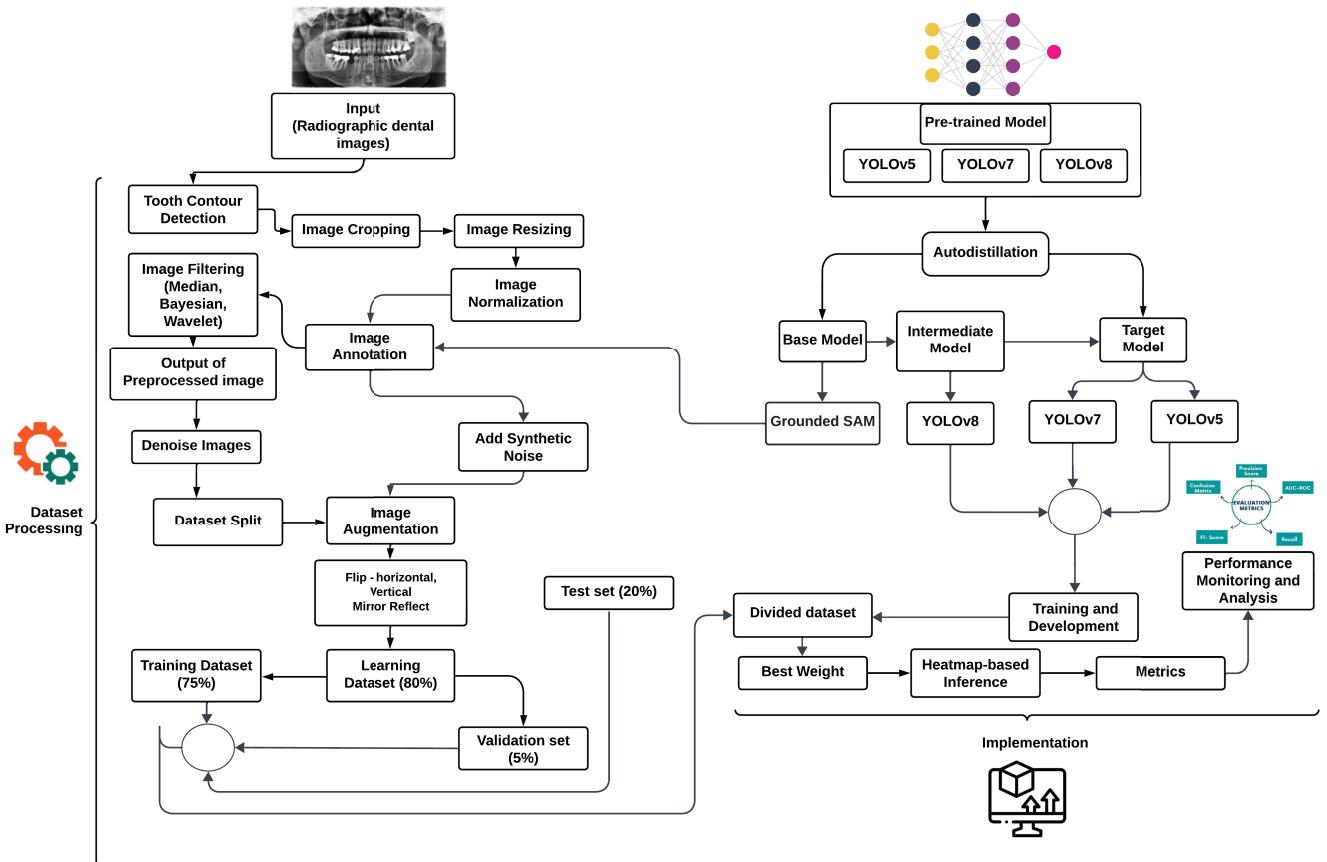


FIGURE 2. Proposed root canal detection system's working progressions.

learning dataset was further split into a training dataset (75%) and a validation dataset (5%). YOLOv5, YOLOv7, and YOLOv8 deep learning models were employed to achieve highly accurate classification and detection. The models' performances were evaluated using mAP metrics, precision curves, and confusion matrix scores. The outputs of the models were then processed with auto-distillation, incorporating the most effective filtering techniques. Grounded SAM served as the base model for generating annotated datasets, while YOLOv8 acted as the intermediate model, and YOLOv5 and YOLOv7 were used as the target models. After auto-distillation, all models were re-evaluated for their respective performances and compared to select the best model for root canal detection. The heatmap-based inference was made by reloading the best weight of the target model to visualize pixels contribution during prediction.

B. DATASET

The subsequent paragraphs describe the creation, preprocessing, and annotation of the dataset utilized in the research.

1) DATASET COLLECTION

Different stages of root canal treatment and their failures are to be detected using a large dataset of radiographic images. For this purpose, approximately 1,600 images were collected

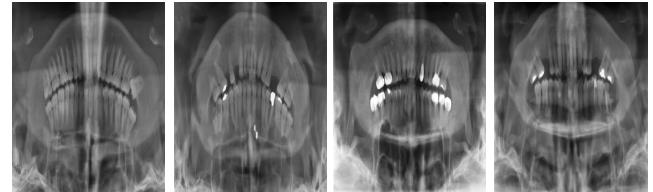


FIGURE 3. Sample images of the custom root canal failure detection dataset.

in May - July 2024 at Square Hospital Ltd, one of the leading medical institutions in Bangladesh, with the assistance of an intraoral X-ray machine. Since only root canal-related images were captured, 1,253 images were selected after filtering out null and blurry images to ensure accurate documentation. This anonymized dataset, showcasing the complex structure of teeth, serves as the baseline for this study. Sample images from our dataset, representing the different stages of root canal treatment, are shown in Figure 3.

2) DATASET CLASSES

This section describes the organization of the images according to the various phases of the root canal treatment process. The dataset is divided into four distinct classes, each representing different stages of root canal treatment and failure, as determined with the assistance of professional dentists [20]. Radiographic images of teeth where the cavity

TABLE 2. Number of training images for each category of the dataset.

Class	Initial images	Augmented images
No Endodontic	405	574
Complete Endodontic	328	660
Incomplete Endodontic	280	700
Total Endodontic Failure	240	760
Total	1253	2694

TABLE 3. Applied augmentation approaches and corresponding parameters.

Augmentation	Parameters
Flip	Horizontal, Vertical
Reflect	Mirror

has reached the pulp but no treatment has been performed are classified as “no endodontic treatment.” When a root canal procedure is in progress, with some infected pulp removed but some remaining for further processing, the images are classified as “incomplete endodontic treatment.” In “complete endodontic treatment,” all infected pulp is removed, and the tooth is sealed with a cap. Some images depict the failure of root canal treatment.

3) DATASET PREPROCESSING

This paragraph outlines the steps taken to clean, normalize, and augment the dataset. After collecting the dataset, data preprocessing was initiated, which included data cleaning, normalization, annotation, and augmentation processes. Initially, all images related to root canal issues were manually sorted, and images of poor quality or with defects were removed. Following this meticulous selection process, the dataset was reduced to 1,253 images, comprising 405 images of “no endodontic treatment,” 328 images of “complete endodontic treatment,” 280 images of “incomplete endodontic treatment,” and 240 images of “total endodontic failure.” The final dataset was divided into a training dataset (80%) and a testing dataset (20%). Table 2 presents the number of training images in each class and the total number of augmented images for each class.

The images with root canal issues were labeled according to their respective classes. The annotation process was carried out using LabelImg, a graphical image annotation tool, by creating bounding boxes around the targeted tooth in each image. After annotation, the images were re-evaluated by dentists to ensure precise labeling. The dataset is also multi-labeled, meaning that specific images may belong to more than one class. This scenario is due to the overlapping clinical features observed by professional dentists.

After annotation, data augmentation is performed to increase the diversity of the dataset using various techniques. The applied augmentation methods and their parameters are detailed in Table 3.

The proposed root canal detection system was implemented and integrated using Azure and Visual Studio. Anaconda is used to manage the packages and Python libraries,

TABLE 4. Tools employed in this work.

Tool	Functions	Purpose
MATLAB	Imgaussfilt(), imfilter(), medfilt2(), imhist()	Image processing and visualization
PyCharm	cv2.GaussianBlur(), cv2.boxFilter(), cv2.medianBlur()	Image processing and visualization
Google Colab Pro	torch.hub.load()	Dataset loading, model training, and performance evaluation

such as PyTorch, for model development, as illustrated in Table 4. Google Colab Pro provided GPU resources, while Roboflow and LabelImg streamlined dataset preparation. Additionally, MATLAB supported image processing and visualization, thereby enhancing the efficiency of the research efforts concerning dental images.

C. APPLIED FILTERING METHODS

In this research, multiple image filtering techniques, i.e., Mean, Median, Contourlet Transform, Bayesian, Non-local, and Gaussian, are applied to improve the employed dataset. Each subsection explains a specific filtering method and its mathematical basis.

1) MEAN

The mean filter is a familiar image processing method used in object detection to reduce noise and smooth images. It replaces each pixel’s value with the average of its neighboring pixels within a specified kernel window. Mathematically, for a kernel of size $m \times n$, the mean filter computes the average intensity at pixel location (x, y) as:

$$M(x, y) = \frac{1}{mn} \sum_{i=-k}^k \sum_{j=-k}^k I(x + i, y + j) \quad (1)$$

where $I(x + i, y + j)$ represents the intensity of the pixel at coordinates $(x + i, y + j)$ in the image.

2) MEDIAN

The median filter is a widely used technique in object detection for noise reduction and preserving edge details. It replaces each pixel’s value with the median value of its neighboring pixels within a defined kernel window. Mathematically, for a kernel of size $m \times n$, the median filter computes the median intensity at pixel location (x, y) as:

$$M(x, y) = \text{median} \left(\begin{array}{c} I(x + i, y + j) \mid i = -k, \dots, k, \\ j = -k, \dots, k \end{array} \right)$$

where $I(x + i, y + j)$ denotes the intensity of the pixel at coordinates $(x + i, y + j)$ in the image.

3) CONTOURLET TRANSFORM

While conventional methods for object detection often rely on wavelet transforms, the Contourlet Transform (CT) offers

distinct advantages due to its superior ability to capture curvilinear image structures. DFB filters utilize directional filters like Gabor filters, which can be expressed as:

$$h(x, y) = g(x, y) \cdot \exp(j2\pi Wx) \quad (2)$$

where $h(x, y)$ is the directional filter, $g(x, y)$ is a windowing function (e.g., Gaussian), and W defines the center frequency in the frequency domain.

4) BAYESIAN

The Bayesian Wavelet denoising algorithm is a sophisticated method for effectively reducing image noise that combines wavelet transform with Bayesian statistical principles. Mathematically, the Bayesian wavelet denoising process adjusts wavelet coefficients $W_{j,k}$ iteratively based on a Bayesian model:

$$\hat{W}_{j,k} = \arg \max_{W_{j,k}} P(W_{j,k} | \text{data}, \sigma^2) \quad (3)$$

where $\hat{W}_{j,k}$ represents the denoised wavelet coefficient at scale j and position k , $P(W_{j,k} | \text{data}, \sigma^2)$ illustrates the posterior probability of $W_{j,k}$ given the data and noise variance σ^2 .

5) NON-LOCAL

The non-local mean (NLM) filter is an advanced image processing technique used in object detection to reduce noise and enhance fine details. This method computes the intensity at each pixel (x, y) as a weighted average of intensities from all pixels in the image, where weights are determined by the similarity between patches centered at (x, y) and other pixels. Mathematically, the non-local mean filter computes the filtered intensity $N(x, y)$ at pixel location (x, y) as:

$$N(x, y) = \frac{1}{C(x, y)} \sum_{p \in \Omega} I(p) \cdot w(p, x, y) \quad (4)$$

where $I(p)$ denotes the intensity of the pixel at position p in the image, Ω is the entire image domain, $w(p, x, y)$ represents the weight based on patch similarity between patches centered at (p) and (x, y) , and $C(x, y)$ is the normalization factor ensuring the sum of weights equals one.

6) GAUSSIAN

Gaussian filtering is a fundamental preprocessing technique in object detection, employing a kernel defined by a Gaussian distribution to convolve with the image. The one-dimensional Gaussian function is expressed as:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (5)$$

This filter reduces noise and preserves critical features by weighting pixel intensities based on their proximity to the kernel center. In two dimensions Gaussian function is expressed as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{(x - \mu)^2 + (y - \mu)^2}{2\sigma^2}\right) \quad (6)$$

σ controls the filter's spread; larger σ values enhance smoothing but may blur edges crucial for object detection, while smaller σ values retain finer details but offer less noise reduction.

Figure 4, Figure 4a, Figure 4b, Figure 4c, Figure 4d, Figure 4e, and Figure 4f demonstrate the original noisy image on the left side and filtered with various filtering methods, i.e., Mean, Median, Contourlet Transform, Bayesian, Non-local, and Gaussian, respectively, on the right side. It is a sample example that shows how the image will transform with its pixel values after the application of filtering.

D. APPLIED DEEP LEARNING MODELS

The following section discusses the applied deep learning models employed for the proposed root canal detection.

1) YOLO

YOLOv5, YOLOv7, and YOLOv8 object detection techniques represent the next generation of object detection models, each with unique design features that enhance efficiency and accuracy [21]. YOLOv5 uses a CSPNet backbone for smooth gradient propagation, a PANet neck to combine multiscale features, and a detection head for bounding box and class predictions, which are optimized during training with loss functions. YOLOv7 incorporates an E-ELAN backbone for improved detection performance and a PANet neck that integrates refined features within a dual-head structure to maximize performance [22]. YOLOv8 extends the previous models by introducing a modified CSPDarknet53 backbone for more efficient feature extraction, a C2F neck for comprehensive feature fusion, and a self-attention mechanism in its head to enhance detection performance further. These models represent significant advancements toward real-time, accurate object detection.

In this work, YOLOv5, YOLOv7, and YOLOv8 were trained on both raw and filtered images. Additionally, individual variants of each model, such as YOLOv5x (extra large), YOLOv5s (small), and YOLOv5l (large), were trained, and only those that delivered the best performance in our experiments were selected.

2) AUTODISTILL MODELS

Autodistill is an optimization technique that utilizes large, slower models to annotate relevant images, which are then distilled into faster, lightweight models [23], [24]. By applying this approach, manual image labeling with a data annotator is no longer required. It involves both a base model and a target model. This study introduces an intermediate model designed to enhance the efficiency of the distillation process. The autodistill technique applied in this study is defined below:

- **Base model** The base model is the initial model in the autodistill process. Labeling relevant images is a time-consuming task that requires specialist data annotators to manually label the images and draw

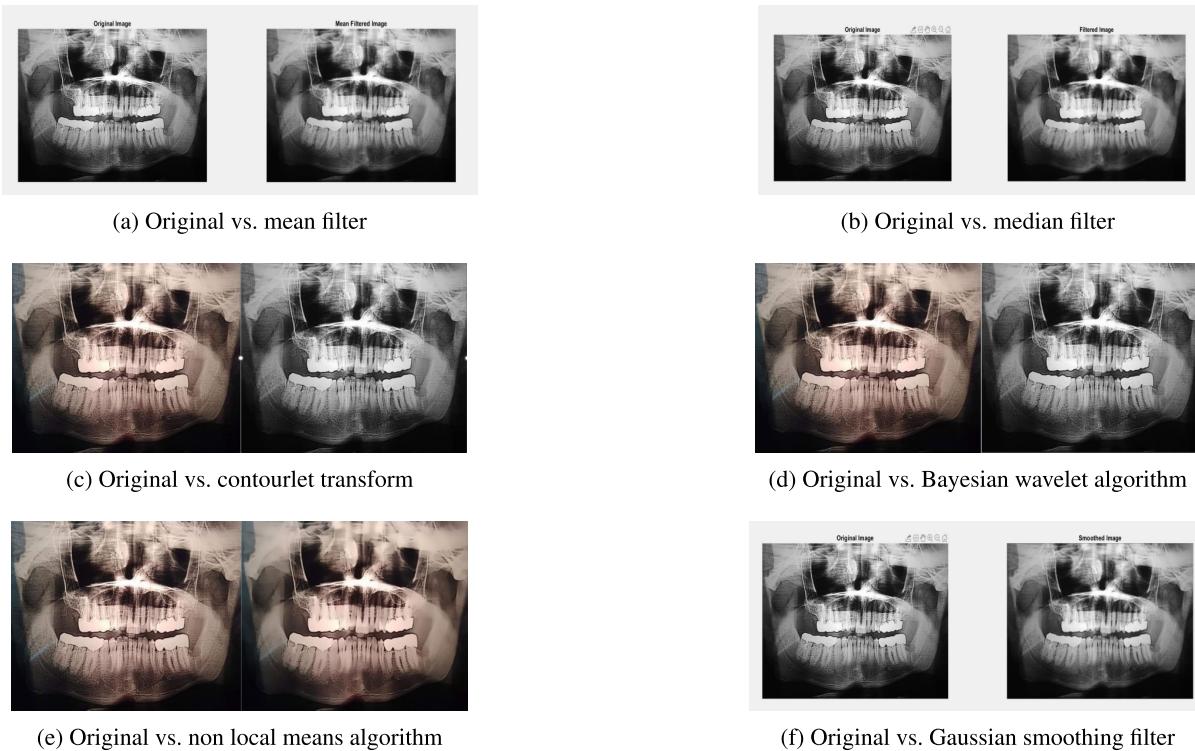


FIGURE 4. Comparison of original and filtered radiographs using various image filtering techniques.

bounding boxes around objects. However, the base model can perform annotation tasks within a day. It must be a large, state-of-the-art (SOTA) model capable of efficiently annotating the required images. The base model processes an Ontology, which provides information on how the model will prompt, the dataset description, and the estimation of the target models. The output of the base model is a newly labeled dataset that can be applied to various tasks, such as detection, segmentation, and others.

- **Intermediate model** The intermediate model is a moderately sized model designed to enhance the efficiency of the distillation process. Its role is to facilitate smoother knowledge distillation, bridge gaps and reduce complexity between the base and target models. It can also act as a form of regularization to help generalize the target models. The intermediate model can increase the target model's training time, leading to faster convergence and improved performance. When a large base model distills knowledge into the intermediate model, it learns and generates a new labeled dataset using its ontology. This process is faster compared to the base model's execution time. The intermediate model consumes the base model-labeled dataset, generates its own labeled dataset based on its knowledge, and can correct errors made by the base model. It can be trained and fine-tuned on the base-labeled dataset or by loading a pre-trained model to generate a new dataset.

- **Target model** Target models operate in the final phase of the autodistill process. These are supervised

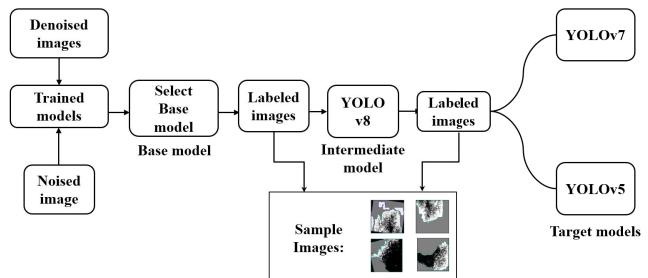


FIGURE 5. Proposed autodistill architecture used in this work.

models that consume the newly labeled dataset. The target model is typically fast, small, and fine-tuned, capable of performing the tasks specified by the newly labeled dataset. After training and evaluating the dataset, it becomes a distilled model. This distilled model is ready for deployment on edge devices. With the distillation knowledge and accurately marked bounding boxes, the model becomes both fast and accurate. The target model is the final output of the autodistill process and can be used to make predictions on unseen images.

Figure 5 illustrates the proposed auto-distillation process with the intermediate model. The process begins with mixed noise and other images, which are then denoised through several filtering methods, such as Bayesian, mean, median, and others. These denoised images are processed by the base models, which generate a new dataset based on their ontology. After evaluation, the dataset, which may contain minor errors, is passed through the intermediate model. The intermediate model, leveraging distilled knowledge, performs

similarly to the base model but generates the dataset more quickly due to its simplicity and better generalization. Finally, the target models are trained on this dataset and generate results. These models become distilled models, ready for deployment on edge devices. The autodistill process offers training and labeling time advantages, smaller model sizes and fewer parameters.

IV. RESULTS

A. EVALUATION OF YOLO MODELS ACROSS VARIOUS FILTERING METHODS

Figures 6, 7, 8, 9, 10, 11, and 12 depicts the confusion matrices of various YOLO models on the all filtered dataset. Figure 6 shows that YOLOv8 demonstrates excellent performance, while YOLOv7 and YOLOv5 show moderate performance. The performance in Figure 7 is similar to that observed for the noisy dataset, with YOLOv8 outperforming the other models. However, it shows a lower performance than with the noisy dataset. Figure 8 indicates that YOLOv5 outperforms the other models despite being smaller and more lightweight. YOLOv5 performs particularly well in the “Complete Endodontic Treatment” class, similar to YOLOv8. Figure 9 shows the confusion matrix for the Contourlet-filtered dataset. YOLOv5 predicts the first two classes (No and Complete Endodontic Treatment) accurately, demonstrating the best performance on the previous datasets. YOLOv8 shows moderate performance, similar to YOLOv7. Figures 10 and 11 present confusion matrices very similar to the Contourlet-filtered dataset, with the Bayesian dataset matrix closely resembling that of the Contourlet-filtered dataset. Finally, Figure 12 shows moderate performance, with all models exhibiting similar performance characteristics.

Table 5 presents the performance of different filtering methods across various YOLO models for the proposed root canal detection system. For raw images, YOLOv5 achieves 78% precision, with recall below 70%, an F1-score of 0.72, and an accuracy of 79.7%. YOLOv7 improves precision to 83%, maintaining almost the same recall as YOLOv5. However, YOLOv8 outperforms both versions across all metrics, achieving a precision of over 90% and an exemplary recall close to 100%, making it significantly more reliable than the other models.

After mean filtering, YOLOv5 achieves a precision of 0.96, an F1-score of 0.88, and an accuracy slightly below 80%. YOLOv7 demonstrates a precision of 0.85, recall of 0.786, and F1-score of 0.816. YOLOv8 achieves a precision and recall of approximately 0.88, with an F1-score of 0.87. While YOLOv5 exhibits the highest precision and F1-score, YOLOv8 shows more moderate performance.

In the case of median filtering, YOLOv5 stands out with the highest scores across all metrics, surpassing 80%. YOLOv7, however, underperforms across all metrics, while YOLOv8 shows excellent accuracy but more moderate scores in other metrics. After applying the Contourlet transform, YOLOv7 achieves a precision of 0.79, recall of 0.74, F1-score of 0.76, and accuracy of 75.0%. In contrast,

YOLOv5 achieves a precision of 0.916, recall of 0.99, F1-score of 0.95, and accuracy of 90.6%, showing excellent performance, similar to the mean-filtered dataset. YOLOv8 also demonstrates high precision and accuracy.

Using the Bayesian wavelet filter, YOLOv5 achieves the highest performance across all metrics (around 90%), indicating strong effectiveness in detecting and correctly identifying objects. YOLOv7, however, exhibits the lowest performance overall. YOLOv8 performs better than YOLOv7 but still falls short of YOLOv5. Like the mean-filtered and Contourlet datasets, YOLOv5 achieves the highest performance among all models on the Bayesian wavelet dataset.

For the Nonlocal Means dataset, YOLOv5 achieves metrics above 80%, though its accuracy is moderate (69.3%). YOLOv7 shows balanced performance across precision, recall, and F1-score, with an accuracy of 90.3%. YOLOv8 stands out with the highest scores across all metrics, demonstrating the most reliable performance.

On the Gaussian filter dataset, YOLOv5 shows a precision of 0.88, recall of 0.866, F1-score of 0.87, and accuracy of 74.0%. YOLOv7 performs slightly lower than YOLOv5, while YOLOv8 shows a precision of 0.84, recall of 0.776, F1-score of 0.796, and accuracy of 80.7%. YOLOv8 has the highest accuracy across all models, while YOLOv5 slightly outperforms it in precision and recall.

The visual outcome in Figure 13 shows predictions at different stages of root canal treatment on both filtered and unfiltered datasets, highlighting the differences in precision, recall, and classification accuracies among the YOLOv5, YOLOv7, and YOLOv8 models. The detected regions are represented by bounding boxes, which indicate the impact of image filtering on model performance. The results demonstrate the comparative strengths of each model in accurately detecting the treatment stages.

B. AUTODISTILLATION TECHNIQUE

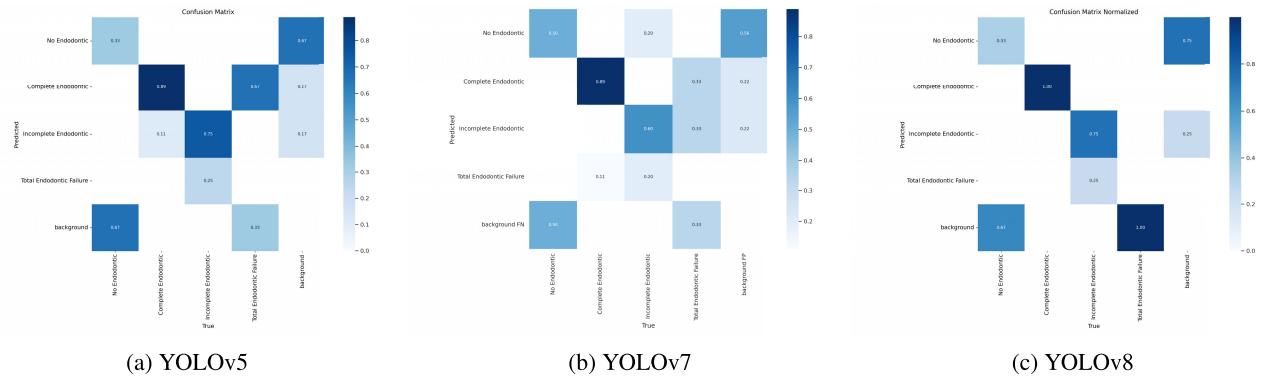
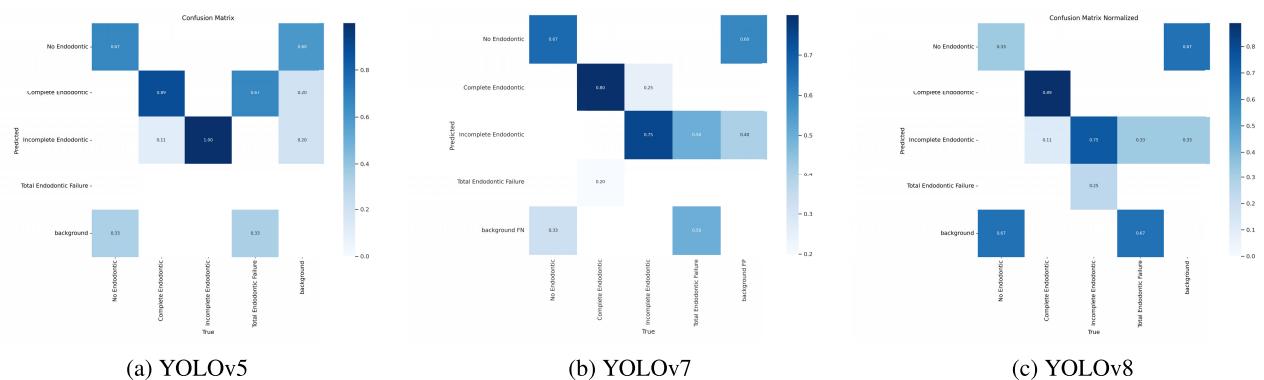
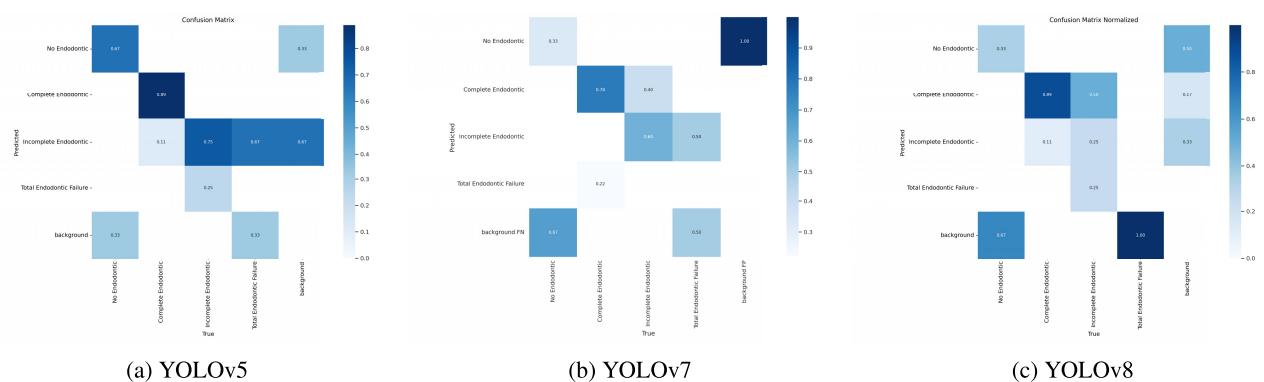
This work also incorporates an optimization technique to further enhance performance. Grounded SAM was selected as the base model to draw bounding boxes for various objects. YOLOv8 serves as the intermediate model, bridging the gap between the base and target models. Given the architectural differences and distinct tasks of Grounded SAM and YOLO models, the intermediate model was introduced to address this issue. Finally, YOLOv5 was chosen over YOLOv7 due to its advantages, including being a lightweight model with fewer parameters and better performance on the filtered dataset.

1) GROUNDED SAM BASE MODEL ANNOTATION

The base model is the initial model used in the autodistillation process. Labeling all images took 34 minutes and 33 seconds, with an average of 3.84 seconds per iteration.

2) YOLOv8 INTERMEDIATE MODEL PERFORMANCE AND PERFORMING REANNOTATION

Figure 14 illustrates the evaluation results for the F1 curve, precision curve, and recall curve. The blue line represents

**FIGURE 6.** Confusion matrix of raw data.**FIGURE 7.** Confusion matrix of mean filtering data.**FIGURE 8.** Confusion matrix of median filtering data.

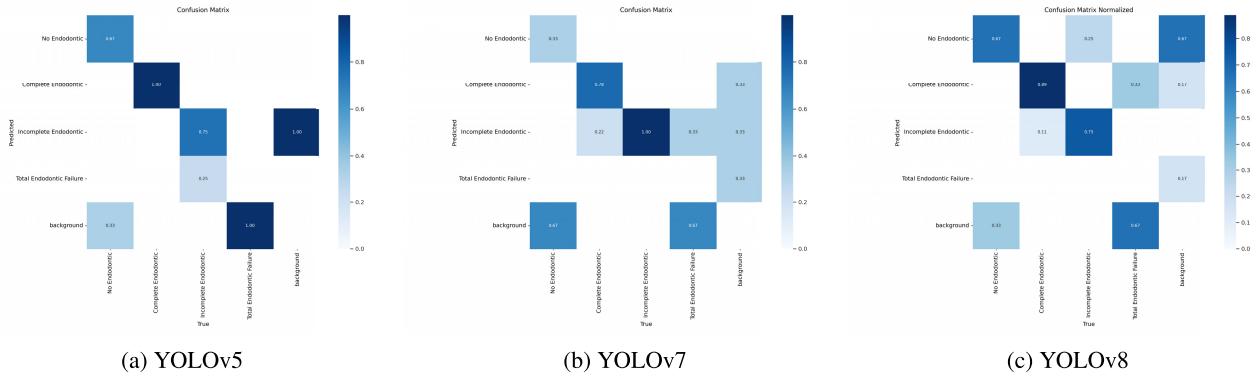
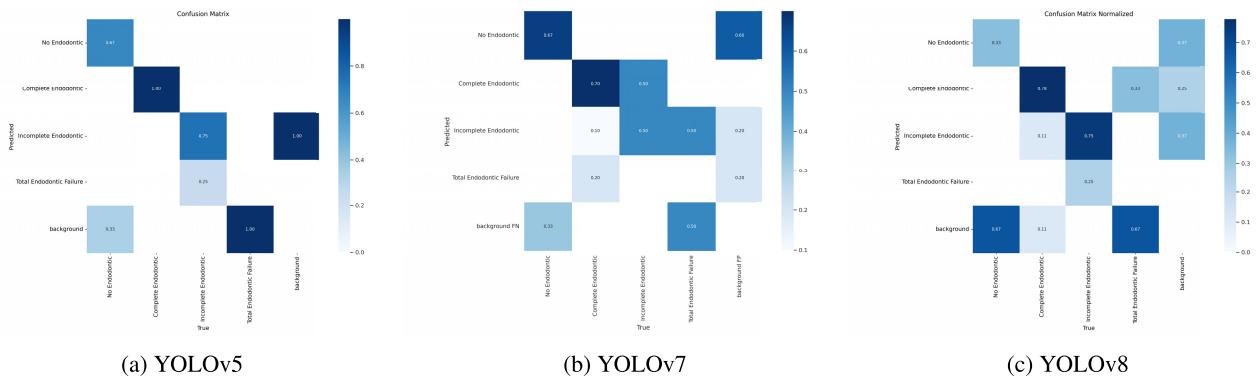
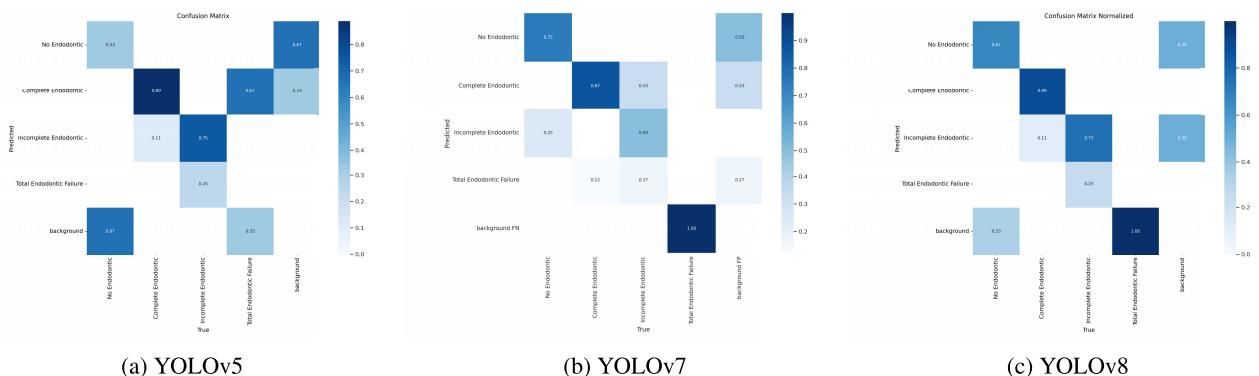
the F1-score, precision, and recall across all classes. The F1-score, illustrated in Figure 14a, peaks at a confidence level of 0.65, reaching approximately 0.76. Precision reaches 1.00 at a confidence level of 0.879, depicted in Figure 14b. The recall is 0.79 at a confidence level of 0.0.

Figure 15 shows the overall performance of the intermediate YOLOv8 model. Training losses seemed to decrease as it progressed, but the validation losses curve was quite different. Precision and Recall seem closer to a straight line after a few epochs passed. The mAP is also quite up and down

when processing a few epochs. After training the intermediate model, the reannotation was executed on the images by loading the pre-trained model. It took only 25 seconds with 21.29 iterations per second, which is very quick compared to the base model.

3) YOLOv5 TARGET MODEL RESULTS

Figure 16 shows the evaluation results for the F1, precision, and recall curves of the YOLOv5 distilled model. The blue line represents the F1-score, precision, and recall across

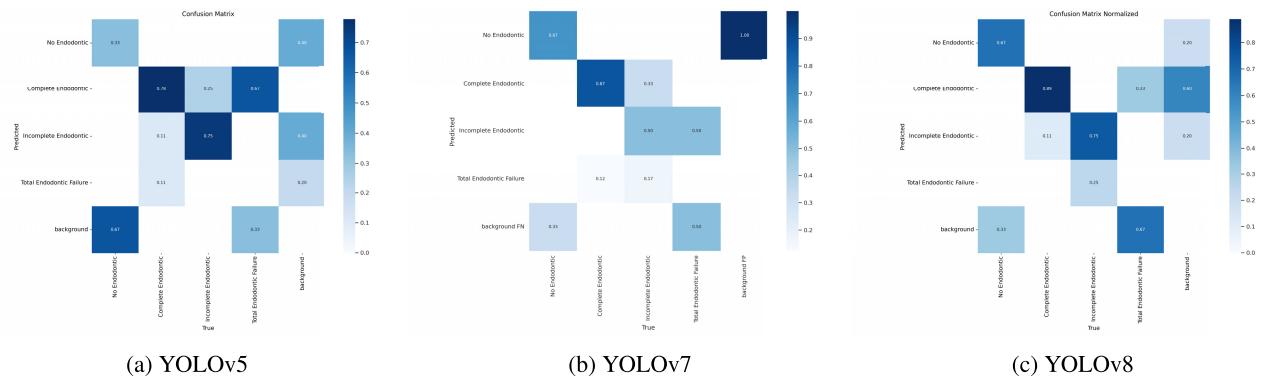
**FIGURE 9.** Confusion matrix of contourlet filtering data.**FIGURE 10.** Confusion matrix of Bayesian filtering data.**FIGURE 11.** Confusion matrix of nonlocal filtering data.

all classes. At a confidence level of 0.57, the F1-score reaches approximately 0.98, close to 1. Precision reaches 1.00 at a confidence level of 0.947, while recall is 1.00 at a confidence level of 0.0. These results demonstrate significant improvements compared to the distilled intermediate models.

Figure 17 shows the overall performance of the distilled YOLOv5 model. Both training and validation losses decreased rapidly as the epochs progressed. Precision and recall exceeded 90, reaching up to 95, demonstrating exceptional performance in the autodistillation process. The

mAP values were also over 80 for IoUs of @50-95 and over 95 for @50 IoU cases. These results indicate substantial improvements in the final stage of the autodistillation process.

Figure 18 presents the results with confidence scores for the distilled YOLOv5 model. Previously, our dataset was classified as multilabeled. The distilled model was applied to classify different treatment stages. It predicted a complete treatment with a high confidence score for all images except one. Additionally, the model made predictions for

**FIGURE 12.** Confusion matrix of Gaussian filtering data.**TABLE 5.** Model performance of the proposed root canal detection system with different filters.

Filter	Model	Precision	Recall	F1-Score	Accuracy
Raw Image	YOLOv5	0.78	0.696	0.72	79.7%
	YOLOv7	0.83	0.696	0.75	69.3%
	YOLOv8	0.92	0.99	0.95	89.3%
Mean Filter	YOLOv5	0.96	0.82	0.88	79.9%
	YOLOv7	0.85	0.786	0.816	77.9%
	YOLOv8	0.88	0.876	0.87	78.2%
Median Filter	YOLOv5	0.88	0.83	0.84	90.6%
	YOLOv7	0.736	0.736	0.76	74.6%
	YOLOv8	0.71	0.776	0.70	89.3%
Contourlet Transform	YOLOv5	0.916	0.99	0.95	90.6%
	YOLOv7	0.79	0.74	0.76	75.0%
	YOLOv8	0.88	0.776	0.816	89.0%
Bayesian Wavelet	YOLOv5	0.916	0.99	0.95	90.6%
	YOLOv7	0.73	0.676	0.706	77.9%
	YOLOv8	0.86	0.866	0.83	88.6%
Non Local Means	YOLOv5	0.88	0.81	0.84	69.3%
	YOLOv7	0.826	0.796	0.74	90.3%
	YOLOv8	0.94	0.956	0.916	89.3%
Gaussian Filter	YOLOv5	0.88	0.866	0.87	74.0%
	YOLOv7	0.79	0.74	0.76	75.0%
	YOLOv8	0.84	0.776	0.796	80.7%

**FIGURE 13.** Predicted results of the applied YOLOv5, YOLOv7 and YOLOv8 models.

incomplete root canal treatments with a low confidence score. These results demonstrate that the autodistillation process successfully compressed the model while maintaining high accuracy.

Heatmap based Inference:

Figure 19 presents heatmap-based inference results from the distilled YOLOv5 model applied to radiographic images, identifying and classifying dental conditions. The bounding boxes and confidence scores indicate the model's detection accuracy for "complete" and "incomplete endodontic treatment."

TABLE 6. Performance of the applied distilled models.

Model	mAP	IoU	Precision	Recall	Accuracy
YOLOv8n (Intermediate)	0.633	@50	0.898	0.651	92.44%
YOLOv5n (Target)	0.994	@50	0.992	0.973	94.09%

Table 6 compares two versions of the YOLOv8n model: Intermediate and Tiny. For instance, the YOLOv8n (Intermediate) model has a mean average precision (mAP) of 63.25, an Intersection over Union (IoU) value of 0.50, a precision rate of 89.84%, a recall value of 65.10%, and an accuracy rate of 92.44%. In contrast, the YOLOv5n (Target) model exhibits a mean average precision (mAP) of 99.40%, which is quite high, with an IoU threshold of 50%. Precision is recorded at 99.16%, while recall is 97.29%, slightly lower but still impressive. Accuracy stands at approximately 94.09%.

These metrics are critical for evaluating model performance, particularly in object detection tasks, as they

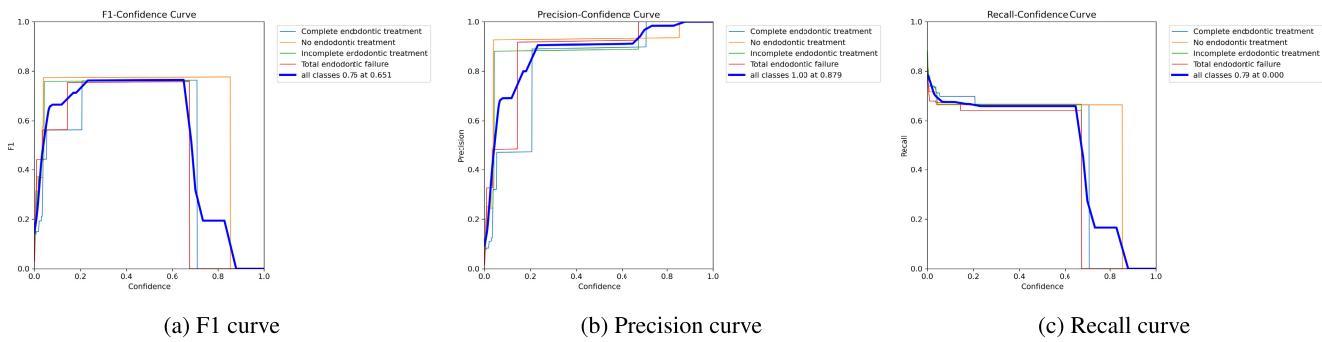


FIGURE 14. Evaluation results of the intermediate YOLOv8 model.

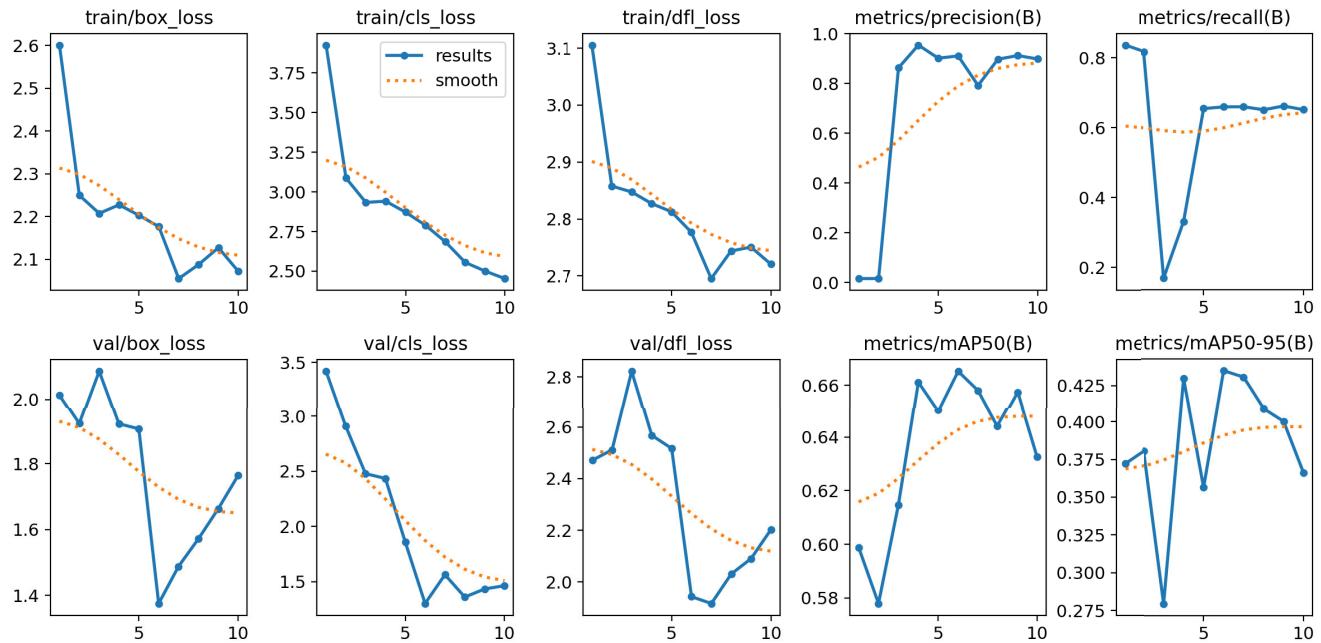


FIGURE 15. Intermediate YOLOv8 model performance and performing re-annotation results.

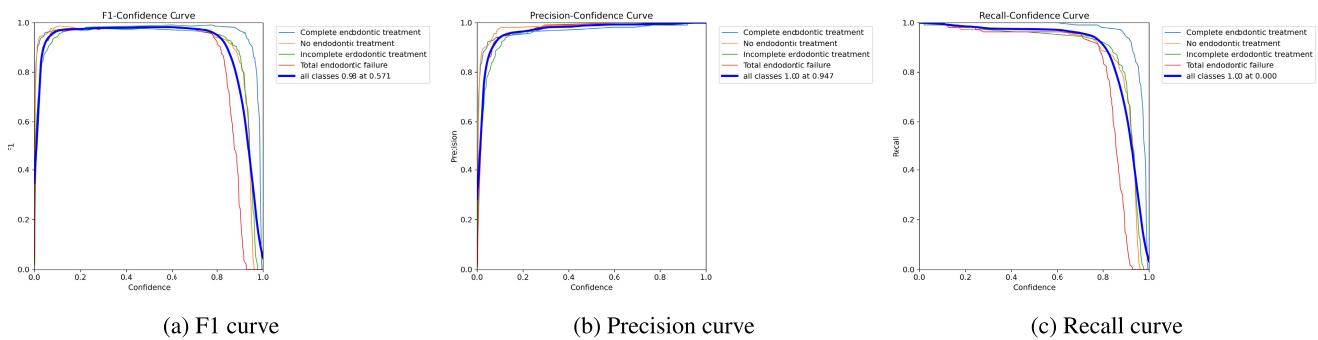


FIGURE 16. Evaluation results of the YOLOv5 target model.

reflect the model's ability to identify objects (precision and recall) correctly. This comparison underscores the significant

improvement from the intermediate to the final form of the YOLOv5n model.

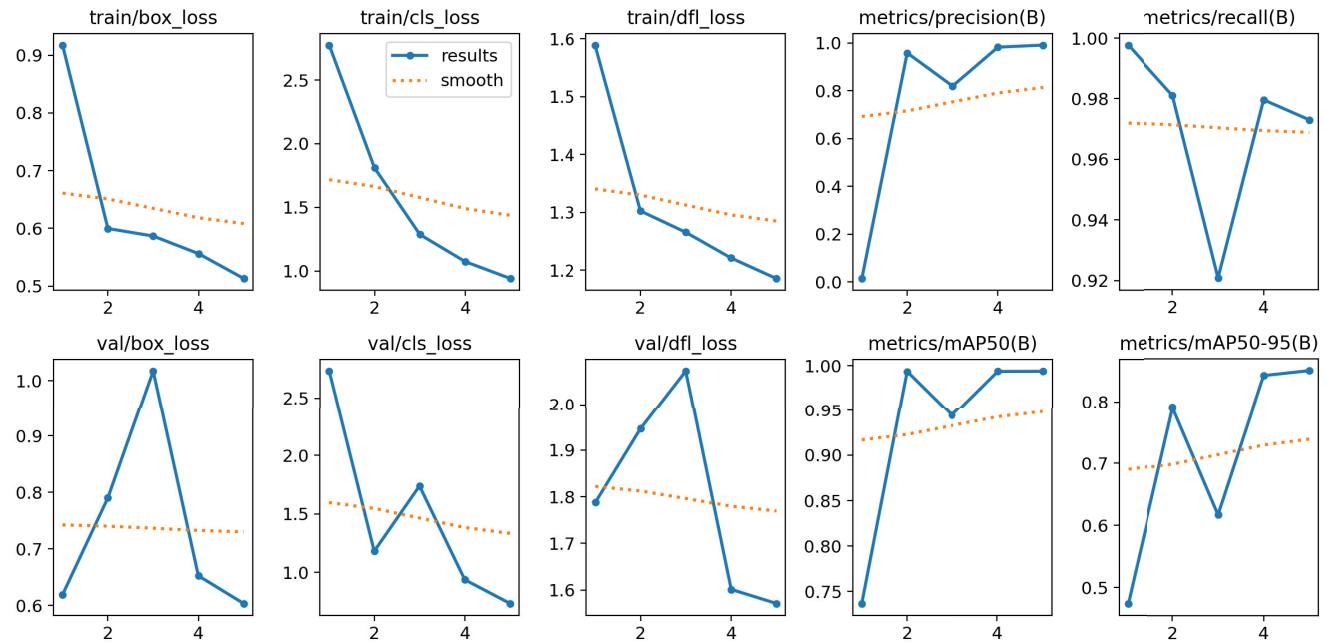


FIGURE 17. YOLOv5 target model performance results.

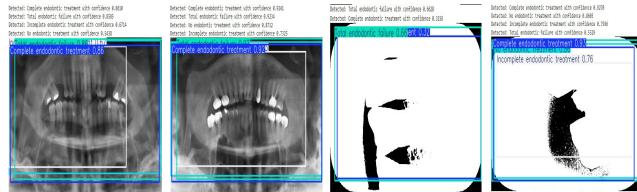


FIGURE 18. Inference results of the distilled YOLOv5 model.

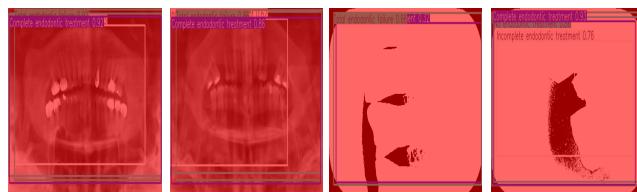


FIGURE 19. Inference results of the distilled YOLOv5 model.

TABLE 7. Comparison of distilled YOLOv5 and YOLOv8 model size and training time per epoch.

Model	Model Size (MB)	Training Time/Epoch (s)
YOLOv8n (Intermediate)	5.96	2.75
YOLOv5n (Target)	5.02	1.94

Table 7 compares distilled YOLOv5 and YOLOv8 in terms of model size and training time per epoch. YOLOv8 features a slightly larger model size (5.96 MB) and requires longer to train per epoch (2.75 seconds) compared to YOLOv5, which is smaller (5.02 MB) and faster (1.94 seconds per epoch).

TABLE 8. Comparison of filtering methods and model accuracy for radiographic detection.

Filtered method	Best Model	Accuracy
Mean Filtering	YOLOv5	79.9%
Median Filtering	YOLOv5	90.6%
Contourlet Transformer	YOLOv5	90.6%
Bayesian Wavelet	YOLOv5	90.6%
Non Local Means	YOLOv7	90.3%
Gaussian Smoothing	YOLOv8	80.7%

Table 8 compares various filtering techniques applied to different YOLO models (YOLOv5, YOLOv7, and YOLOv8) in terms of accuracy. The methods considered include mean filtering, median filtering, contourlet transform, Bayesian wavelet, nonlocal filtering, and Gaussian smoothing. The accuracy percentages range from 79.9% for mean filtering on YOLOv5 to 90.6% for median filtering, contourlet transform, and Bayesian wavelet on YOLOv5. Additionally, nonlocal filtering achieved 90.3% accuracy on YOLOv7, while Gaussian smoothing resulted in the lowest accuracy of 80.7% on YOLOv8. This comparison highlights how different object detection models perform with various filtering techniques, some significantly improving accuracy.

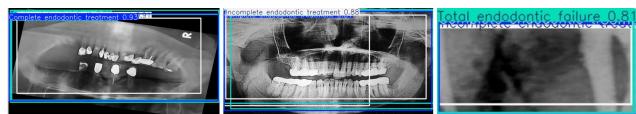
Table 9 shows the accuracy of different filtering techniques applied to a raw dataset using various YOLO models. It also highlights the performance improvements achieved with different variants of the auto-distillation model. The raw dataset has an accuracy of 89.3% when processed with the YOLOv8 model. In contrast, the filtered dataset achieved better performance with three filtering techniques: median

TABLE 9. Accuracy comparison of YOLO models on Raw, filtered, and distilled datasets.

Dataset	Description	Best Model	Accuracy (%)
Raw Dataset	Manual Filtering	YOLOv8	89.3
	Median Filtering	YOLOv5	90.6
Filtered Dataset	Contourlet Transformer	YOLOv5	90.6
	Bayesian Wavelet	YOLOv5	90.6
Auto-Distilled Models	YOLOv8n (Intermediate)	YOLOv8n	92.44
	YOLOv5n (Target)	YOLOv5n	94.09

TABLE 10. Resource and model specifications with inference speed.

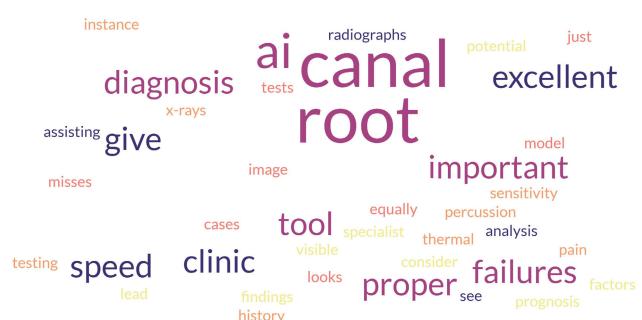
Resource / Model	Specification / Inference Speed
GPU	Google Colab Pro (40GB VRAM) and Kaggle GPU
Batch Size	16
Learning Rate	0.001 with cosine annealing scheduler
Optimizer	SGD with momentum 0.937 and weight decay 0.0005
Training Epochs	50
Image Size	640×640
YOLOv8 Inference Speed	24ms per image
Knowledge-distilled YOLOv5 Inference Speed	12ms per image
Preprocessing (filtering) Time	8ms per image
Total Processing Time	~20ms per image

**FIGURE 20.** Prediction results of the YOLOv5n model tested on the dataset Provided by Hasan et al. [25].

filtering, Contourlet transform, and Bayesian wavelet, each reaching 90.6% accuracy. Evaluation of the intermediate and target models of auto-distillation on YOLOv8n and YOLOv5n resulted in accuracy values of 92.44% and 94.09%, respectively. For YOLOv8, the application of the intermediate model improved accuracy by 3.14%, while the target model provided an additional 3.49% increase. These results demonstrate that filtering methods enhance the performance of object detection models while auto-distillation further improves their accuracy.

Table 10 summarizes the utilized hardware resources, training parameters, and inference performance of the distilled YOLOv5n model used in this study. The knowledge-distilled YOLOv5 model demonstrated better computational efficiency, achieving an inference speed of 12 ms per image, i.e., twice as fast as YOLOv8 (24 ms per image). Including preprocessing (8 ms), the total per-image processing time is approximately 20 ms. Considering real-time systems typically require processing speeds under 33 ms per frame to meet 30 FPS standards, the proposed model demonstrates strong feasibility for near real-time clinical deployment.

The distilled YOLOv5n model with median filtering technique was employed to test the dataset provided by Hasan et al. [25], identifying bounding boxes and confidence scores for distinct stages of endodontic treatment. The distilled YOLOv5n model attained 86.94% accuracy for the 240 test samples of the dataset, demonstrating its generalizability. As shown in Figure 20, the model successfully classified various critical conditions in dental radiographs:

**FIGURE 21.** Word cloud of expert feedback on the proposed root canal failure diagnosis system.

Complete Endodontic Treatment, Incomplete Endodontic Treatment and Total Endodontic Failure. Complete Endodontic Treatment was detected with high confidence scores, underscoring the model's reliability in recognizing well-executed procedures. Peak confidence scores of 0.9908 and 0.9695 further emphasized the robustness of these detections. Incomplete Endodontic Treatment was in multiple radiographs, with confidence scores between 0.28 and 0.89. While an anomalous score of 8.8872 was recorded (likely an outlier), more representative detections included scores of 0.7201 and 0.3601, indicating procedural deficiencies that may necessitate clinical follow-up. Total Endodontic Failure was in several cases, with confidence scores spanning from 0.4145 to 0.9293. The highest confidence score of 0.9293 highlighted severe complications requiring urgent intervention, while additional detections (e.g., 0.9010 and 0.8847) validated the model's efficacy in identifying failed treatments. As illustrated in Figure 20, the bounding boxes visually delineate these treatment stages, highlighting the model's capacity for nuanced classification of endodontic outcomes.

A pilot study was conducted in collaboration with two dental professionals to validate the practical effectiveness of the proposed model. Five representative radiographic images from each of the four classes of root canal treatment and failure were presented to the experts before and after applying the YOLOv5n detection model. The dental specialists evaluated the AI model's predictions objectively, discussing intuition about the proposed root canal failure diagnosis system. Figure 21 demonstrates the word cloud summarizing the obtained expert feedback.

TABLE 11. Comparison of the proposed root canal detection system with similar works.

Ref.	Dataset	Method	Best Model	Metrics
[10]	Images: 435	Two stage neural network: object detection and binary classification	CNN	Accuracy: 93.4%
[11]	Images: 26316	Two-step approach: DC-ChestNet and VT-ChestNet	VT-ChestNet	AUC: 99.57%
[16]	Images: 636	Pre-trained deep learning models	ResNet50	Accuracy: 72.00%
[17]	Images: 1352 (adults)	Deep learning CNN model for teeth detection and numbering	VGG16	F1-score: 99.80%
[18]	Images: 915	LSTM and CNN deep learning methods	Filtered-CNN	Accuracy: 84.37%
[7]	Images: 4821	YOLOv4 and CNN deep learning model based on object detection and classification	YOLOv4	F1-score: 95.00%
This work	Images: 1600	Denoising methods and YOLOv5, YOLOv7, and YOLOv8 models, applying auto-distillation to the best models	Filtered YOLOv5 distilled model Unfiltered YOLOv8 distilled model	Accuracy: 94.1% Accuracy: 92.4%

Table 11 illustrates a comparison of the proposed root canal detection system with similar works. The proposed system uses advanced versions of the YOLO algorithm and further filtering models and auto-distillation techniques as described above, hence giving it better identification capabilities than previous methods and making it more preferred for endodontic detection.

V. DISCUSSION

This study suggested deep learning-based diagnosing and detecting failures of root canals based on neural networks and sophisticated image analysis tools. Automated and quantitative medical evaluations using the proposed neural network-integrated system in place of subjective evaluations by dental professionals to human error and discrepancies in treatment outcomes and, consequently, quality of dental care are expected.

The study regards data quality as necessary to build robust predictive models. The images were denoised using pre-processing techniques such as Gaussian smoothing followed by median and mean filtering, which improved the input data quality. However, care should always be taken not to discard too much diagnostic information by denoising. Future studies should develop better techniques regarding detail preservation with noise reduction.

Conduction of data preparation into training, validation and test splits was well per standard; a breakdown into subset composition and failure types would impel detailed capital reproduction of the results. Such detail would lend itself to the improvement of the model's performance and generalization variability for heterogeneous cases.

The application of YOLOv5 for the training and evaluating models indicates that the work is within the modern deep learning feature extraction frameworks. Though the performance has shown a precision of 89.1% and mAP of 0.636, the performance validation should be across diverse datasets and clinical settings. The distilled YOLOv5n model was evaluated on both our collected dataset and an independent dataset. Future improvements should include expansion of the dataset, addition of clinical parameters, and validation studies, all of which are for the ultimate goal of clinical applicability of the system.

VI. IMPACTS OF THE RESEARCH

The developed system promises to change the face of endodontic diagnosis and treatment. Rapid and accurate detection of root canal problems will likely result in less time and money spent on complicated oral treatments. Early diagnosis, made possible by the system, allows the patient to start treatment earlier when outcomes are better, and resource use is lower.

This work also encourages the adoption of readily available digital radiographic imaging and electronic health records. It will eventually reduce reliance on traditional film-based radiographs and paper documentation toward resource efficiency and sustainability, involving lower waste.

The predictive model can even allow dentists to customize the treatment schedules by tailoring them to various individual characteristics affecting efficacy in root canal procedures. Furthermore, it inspires proactive prevention in terms of oral healthcare that involves regular visits, early intervention, and maintenance of overall oral health.

For successful implementation, collaboration with dentists, legal compliance, and consideration of sociocultural factors would be critical in maximizing the positive impacts of this work while addressing the concerns regarding potential barriers.

VII. CONCLUSION

The viability of employing deep learning models and de-noising algorithms toward enhancing the detection of root canal failures is proved and quite well followed in this study. Applying YOLOv5, YOLOv7, and YOLOv8 models to the radiographic images acquired from the local clinics has recorded promising results, with YOLOv5 even producing the highest accuracy. There was increased performance, even using knowledge distillation, making a case for how compelling techniques could be combined for object detection into a single setting. The proposed work in this research utilizes denoising methods and YOLOv5, YOLOv7, and YOLOv8 models with auto-distillation to achieve an accuracy of 92.4% on a dataset of 1600 images. However, the study comes with limitations. This study did not consider other forms of dental treatment but root canal treatment alone; additionally, the dataset was only region-specific.

Future research needs to be obtained from various data sources and multiple dental institutions within different geographies, and further models need to be developed to classify different dental diseases. Such improvements may include fine-tuning training data, applying ensemble models for annotation accuracy, or adding explainable AI techniques, e.g., Class Activation Mapping (CAM) or Grad-CAM, to achieve better results. Moreover, real-time deployment on cloud platforms such as Google Cloud or AWS can open up more access and practical implementation. These challenges could recur while developing research to succeed in this current study. These advancements are critical to revolutionizing endodontic care and improving patient outcomes.

REFERENCES

- [1] M. T. Yactayo-Albuquerque, M. L. Alen-Méndez, D. Azañedo, D. Comandé, and A. Hernández-Vásquez, "Impact of oral diseases on oral health-related quality of life: A systematic review of studies conducted in Latin America and the Caribbean," *PLoS ONE*, vol. 16, no. 6, Jun. 2021, Art. no. e0252578.
- [2] R. Jurič, G. Vidmar, R. Blagus, and J. Jan, "Factors associated with the outcome of root canal treatment—A cohort study conducted in a private practice," *Int. Endodontic J.*, vol. 57, no. 4, pp. 377–393, Apr. 2024. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/iej.14022>
- [3] J. O. Makanjuola, O. H. Oderinu, and D. C. Umesi, "Treatment outcome and root canal preparation techniques: 5-year follow-up," *Int. Dental J.*, vol. 72, no. 6, pp. 811–818, Dec. 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0020653922002015>
- [4] L. E. Burns, J. Kim, Y. Wu, R. Alzwaideh, R. McGowan, and A. Sigurdsson, "Outcomes of primary root canal therapy: An updated systematic review of longitudinal clinical studies published between 2003 and 2020," *Int. Endodontic J.*, vol. 55, no. 7, pp. 714–731, Jul. 2022. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/iej.13736>
- [5] C. P. Behrenbruch, S. Petroudi, S. Bond, J. D. Declerck, F. J. Leong, and J. M. Brady, "Image filtering techniques for medical image post-processing: An overview," *Brit. J. Radiol.*, vol. 77, no. 2, pp. S126–S132, Dec. 2004.
- [6] C. Yi Li, R. Mazzon, and A. Cavallaro, "Underwater image filtering: Methods, datasets and evaluation," 2020, *arXiv:2012.12258*.
- [7] E. Kaya, H. G. Güneç, E. Ş. Ürkmez, and K. C. Aydün, "Deep learning for diagnostic charting on pediatric panoramic radiographs," *Int. J. Computerized Dentistry*, Jul. 2023.
- [8] K.-C. Li, Y.-C. Mao, M.-F. Lin, Y.-Q. Li, C.-A. Chen, T.-Y. Chen, and P. A. R. Abu, "Detection of tooth position by YOLOv4 and various dental problems based on CNN with bitewing radiograph," *IEEE Access*, vol. 12, pp. 11822–11835, 2024.
- [9] L. Chen, J. He, Y. Wu, Y. Tang, G. Ge, and W. Wang, "Detection and 3-D visualization of human tooth surface cracks using line structured light," *IEEE Sensors J.*, vol. 24, no. 9, pp. 13958–13967, May 2024.
- [10] B. Pridgen, L. von Rabenau, A. Luan, A. J. Gu, D. S. Wang, C. Langlotz, J. Chang, and B. Do, "Automatic detection of perilunate and lunate dislocations on wrist radiographs using deep learning," *Plastic Reconstructive Surg.*, vol. 153, no. 6, pp. 1138e–1141e, Jun. 2024.
- [11] A. A. Nasser and M. A. Akhloufi, "Deep learning methods for chest disease detection using radiography images," *Social Netw. Comput. Sci.*, vol. 4, no. 4, May 2023.
- [12] M. Javed, M. Zahra, and M. Ali, "Use of deep neural networks with radiographic images for automated COVID-19 detection," *Biol. Clin. Sci. Res. J.*, vol. 2023, no. 1, p. 249, Apr. 2023.
- [13] M. Mann, R. P. Badoni, H. Soni, M. Al-Shehri, A. C. Kaushik, and D.-Q. Wei, "Utilization of deep convolutional neural networks for accurate chest X-ray diagnosis and disease detection," *Interdiscipl. Sci., Comput. Life Sci.*, vol. 15, no. 3, pp. 374–392, Sep. 2023.
- [14] W. Hou, D. Zhang, Y. Wei, J. Guo, and X. Zhang, "Review on computer aided weld defect detection from radiography images," *Appl. Sci.*, vol. 10, no. 5, p. 1878, Mar. 2020. [Online]. Available: <https://www.mdpi.com/2076-3417/10/5/1878>
- [15] W. Shi, L. Tong, Y. Zhu, and M. D. Wang, "COVID-19 automatic diagnosis with radiographic imaging: Explainable attention transfer deep neural networks," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 7, pp. 2376–2387, Jul. 2021.
- [16] B. Çelik and M. E. Çelik, "Root dilaceration using deep learning: A diagnostic approach," *Appl. Sci.*, vol. 13, no. 14, p. 8260, Jul. 2023.
- [17] D. V. Tuzoff, L. N. Tuzova, M. M. Bornstein, A. S. Krasnov, M. A. Kharchenko, S. I. Nikolenko, M. M. Sveshnikov, and G. B. Bednenko, "Tooth detection and numbering in panoramic radiographs using convolutional neural networks," *Dentomaxillofacial Radiol.*, vol. 48, no. 4, May 2019, Art. no. 20180051.
- [18] C. Buyuk, B. Arican Alpay, and F. Er, "Detection of the separated root canal instrument on panoramic radiograph: A comparison of LSTM and CNN deep learning methods," *Dentomaxillofacial Radiol.*, vol. 52, no. 3, Feb. 2023.
- [19] H. Karkehabadi, E. Khoshbin, N. Ghasemi, A. Mahavi, H. Mohammad-Rahimi, and S. Sadr, "Deep learning for determining the difficulty of endodontic treatment: A pilot study," *BMC Oral Health*, vol. 24, no. 1, May 2024.
- [20] A. Tareq, M. I. Faisal, M. S. Islam, N. S. Rafa, T. Chowdhury, S. Ahmed, T. H. Farook, N. Mohammed, and J. Dudley, "Visual diagnostics (don't short) of dental caries through deep learning of non-standardised photographs using a hybrid YOLO ensemble and transfer learning model," *Int. J. Environ. Res. Public Health*, vol. 20, no. 7, p. 5351, Mar. 2023.
- [21] J. Terven, D.-M. Córdoba-Esparza, and J.-A. Romero-González, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Mach. Learn. Knowl. Extraction*, vol. 5, no. 4, pp. 1680–1716, Nov. 2023. [Online]. Available: <https://www.mdpi.com/2504-4990/5/4/83>
- [22] C.-Y. Wang, A. Bochkovskiy, and H.-Y.-M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 7464–7475.
- [23] M. F. S. Titu, M. A. Pavel, G. K. O. Michael, H. Babar, U. Aman, and R. Khan, "Real-time fire detection: Integrating lightweight deep learning models on drones with edge computing," *Drones*, vol. 8, no. 9, p. 483, Sep. 2024.
- [24] X. Zhang, Z. Zhou, D. Chen, and Y. Emma Wang, "AutoDistill: An end-to-end framework to explore and distill hardware-efficient language models," 2022, *arXiv:2201.08539*.
- [25] H. A. Hasan, F. H. Saad, S. Ahmed, N. Mohammed, T. H. Farook, and J. Dudley, "Experimental validation of computer-vision methods for the successful detection of endodontic treatment obturation and progression from noisy radiographs," *Oral Radiol.*, vol. 39, no. 4, pp. 683–698, Oct. 2023.



AFIFA ZAIN APURBA is currently an Energetic Undergraduate with the Department of Electrical and Computer Engineering, North South University, Bangladesh. She has a keen interest in artificial intelligence and its new applications. Her research ranges in frontier areas from vision-language models to object detection, with further involvement in machine learning. She has done considerable service to academia, such as "Revolutionising Workout Analytics: Utilizing Machine Learning Models for Estimation of Calories Burnt" and studying pedagogic strategies for the food processing industry. Her passion for finding solutions through AI speaks volumes about what she could achieve as a future leader in technology and research.



MD FAHIM SHAHORIAR TITU received the B.Sc. degree in computer science and engineering (CSE) from the Department of Electrical and Computer Engineering, North South University, Bangladesh, in 2023. He is a Talented Data Analyst and a Dedicated Researcher with a strong foundation in data science, robotics, and deep learning. He is currently a Data Analyst with OpenFirm, he leverages his expertise to tackle complex and real-world problems. His impactful research has been published in prestigious journals, such as *Drones*, *Sci*, and *Asian Journal of Medical Technology*. He is passionate about innovation and enjoys bridging the gap between advanced computational techniques and practical applications, working on cutting-edge solutions including, but not limited to, deep learning-based visual pollution detection, real-time fire detection via drone usage, and cooperative robotics for industrial environments. He believes in exploring creative intersections between AI, robotics and data analytics, and representing a continued impact on technology.



MAHIR AFSER PAVEL received the B.Sc. degree in computer science and engineering from North South University, Dhaka, Bangladesh. He is currently pursuing the Internship and an Artificial Intelligence Engineer with EkkBaz. His research interests include natural language processing, machine translation, medical image analysis, computer vision, and related subfields.



INTISAR TAHMID NAHEEN received the B.Sc. degree in electrical and electronics engineering from the Islamic University of Technology, Gazipur, Bangladesh, in 2014, and the M.Sc. degree in information and communication engineering from Technische Universität Darmstadt (TU Darmstadt), Germany, in 2018. Currently, he is a Senior Lecturer with the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh. His current research interests include artificial intelligence (AI), reinforcement learning (RL), computer vision, 3D object detection, augmented reality, and application of machine learning in the power sector.



RIASAT KHAN (Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from the Islamic University of Technology, Bangladesh, in 2010, and the M.Sc. and Ph.D. degrees in electrical engineering from New Mexico State University, Las Cruces, USA, in 2018. Currently, he holds the position of an Associate Professor with the Department of Electrical and Computer Engineering, North South University, Dhaka, Bangladesh. His research interests include data science, machine learning, computational bioelectromagnetics, and power electronics.

• • •