



OPEN ACCESS

EDITED BY

Viresh Chopra,
Oman Dental College, Oman

REVIEWED BY

Surendra Maharjan,
NewYork-Presbyterian, United States
Hadeel Mazin Akram,
University of Baghdad, Iraq

*CORRESPONDENCE

Iman AlMomani

✉ i.momani@ju.edu.jo

RECEIVED 22 October 2025

REVISED 26 January 2026

ACCEPTED 02 February 2026

PUBLISHED 27 February 2026

CITATION

Aljamani S, AlMomani I, El-Shafai W, AL-Akhras M, AlHaddad A and Abu zaghlan R (2026) Developing an AI-powered tool for radiographic feedback on working length determination in pre-clinical endodontic training.
Front. Dent. Med. 7:1730454.
doi: 10.3389/fdmed.2026.1730454

COPYRIGHT

© 2026 Aljamani, AlMomani, El-Shafai, AL-Akhras, AlHaddad and Abu zaghlan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Developing an AI-powered tool for radiographic feedback on working length determination in pre-clinical endodontic training

Sanaa Aljamani^{1,2}, Iman AlMomani^{3,4,5*}, Walid El-Shafai^{6,7}, Mousa AL-Akhras⁸, AbdulAziz AlHaddad¹ and Rawan Abu zaghlan^{1,2}

¹Restorative Department, School of Dentistry, The University of Jordan, Amman, Jordan, ²Consultant in Endodontics, Restorative Department, Jordan University Hospital, Amman, Jordan, ³Computer Science Department, King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan, ⁴Intelligent Cybersecurity Engineering Research Group, The University of Jordan, Amman, Jordan, ⁵Computer Science Department, CCIS, Prince Sultan University, Riyadh, Saudi Arabia, ⁶Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, ⁷Automated Systems and Computing Lab (ASCL), Prince Sultan University, Riyadh, Saudi Arabia, ⁸Computer Information Systems Department, King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan

Background: Establishing an accurate working length is a critical step in root canal treatment and directly influences clinical success. As artificial intelligence increasingly integrates into medical education, applying it to enhance endodontic training has become increasingly important.

Aims: This study aimed to develop a machine learning-based tool that provides prompt, personalized, constructive feedback on radiographic working length determination in a pre-clinical setting and to evaluate its usability among dental students.

Methods: A newly labeled dataset of 3,000 radiographic images was created and categorized into optimal, over-extended, and under-extended working lengths. This dataset was balanced and split into 80%, 10%, and 10% for training, validation, and testing, respectively. Twenty-two convolutional neural network models were developed, trained, and evaluated using five diagnostic metrics (accuracy, F1-score, precision, recall, and testing time). The best-performing model was integrated into a web-based platform and piloted with 30 pre-clinical dental students who provided usability feedback via a Likert-scale questionnaire. The study hypothesized that students would rate the tool as usable and educationally supportive.

Results: The custom-developed deep CNN achieved 97%–99% accuracy, 95%–98% F1-score, 94%–99% precision, and a recall rate of 96%–98%, with an average testing time of 0.54 s. Students rated the proposed system positively across clarity, ease of use, and learning support, with median usability scores of 5.0 across all items and interquartile ranges of 4–5 to 5–5.

Conclusion: The AI-powered feedback system demonstrated high accuracy with strong user acceptance. By delivering instant, constructive feedback on working length determination, it supports effective learning and skill refinement in endodontic education. It is also beneficial in classrooms with large student populations. Future work will expand the dataset and integrate additional stages of root canal training into a unified AI-based educational platform.

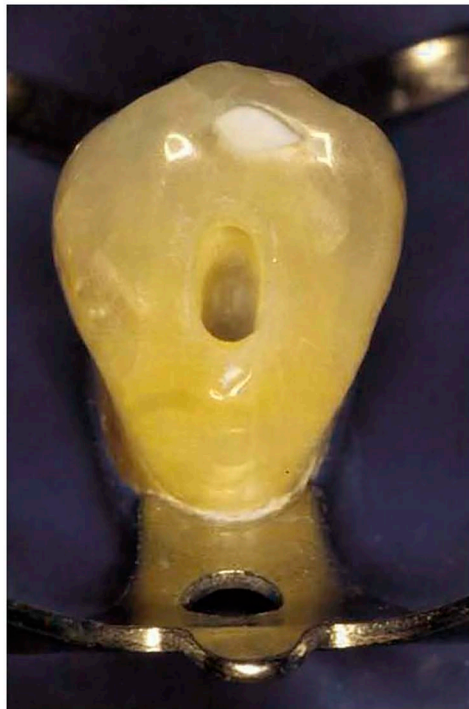
KEYWORDS

artificial intelligence, constructive feedback, convolutional neural network (CNN), education technology, endodontic education, ground-truthing, machine learning, radiographic

1 Introduction

Root canal treatment (RCT) is a fundamental procedure in endodontics aimed at eliminating dental pulp infections through the cleaning and shaping of the root canal system,

utilizing specialized instruments and materials (1, 2). Figure 1 illustrates different stages of RCT, which involve chemical and mechanical cleaning to the correct working length of the root canal and filling the space with a dedicated root canal-filling material.



(a) Access cavity preparation.



(b) Working length determination.



(c) Master cone fit.



(d) Obturation of the root canal.

FIGURE 1
The technical stages of root canal treatment.

The success of RCT is influenced by various factors, including tooth-related aspects, accurate diagnosis, and careful interpretation of radiographic examinations (3, 4). The technical stages of RCT follow a sequential approach, beginning with access cavity preparation, followed by working length (WL) determination, chemo-mechanical preparation, and obturation. Each step is codependent, with the accuracy of one step directly affecting the next (1, 4, 5).

Working length (WL) in RCT defines the extent to which cleaning and shaping should be performed within the root canal (4). Literature suggests that WL should terminate within 2 mm of the radiographic apex to optimize treatment outcomes (6–8). Incorrect working length determination can lead to undesirable technical and biological consequences of the whole endodontic procedure (6, 9, 10). Historically, WL has been determined using various techniques, including patient response and tactile sensation. However, these techniques are often inaccurate due to factors such as root canal morphology, apical inflammation, and patient variability (11, 12). Electronic apex locators' technique has demonstrated improved accuracy in WL determination and reduced radiation exposure, but is also associated with limitations when used alone (8, 13, 14). A combined approach incorporating radiographic WL measurements is recommended to ensure reliable results of WL determination (7, 15, 16).

Artificial intelligence (AI) is increasingly being adopted in the dental field to enhance efficiency and accuracy (17–20). AI-driven machine learning models based on radiographic and image analysis have shown promising results in diagnosing dental caries (21), detecting periapical pathology and diagnosis (19, 22–24), assessing the quality of root canal filling (25, 26), and improving orthodontic treatment planning (27, 28).

The research on AI applications in WL determination remains limited (28). In one study, AI-driven models for WL determination have demonstrated considerable accuracy, with some models achieving 96% accuracy compared to specialist endodontists 76% (29). Another model achieved 85% accuracy compared to the dual-frequency impedance ratio method (30). Table 1 summarizes relevant studies that explored AI work in radiographic examination in endodontics.

Undergraduate dental education introduces students to the science and practice of endodontics through pre-clinical training, where they develop knowledge, skills, and confidence via theoretical and practical sessions (31–34). These sessions emphasize the importance of key factors influencing successful RCT, such as understanding tooth anatomy, accurate diagnosis, careful interpretation of radiographs, and practical steps of root canal treatment (3, 35, 36).

While AI integration in dental education has been generally well-received, such as virtual reality simulators for operative dentistry (37–39) and AI-assisted diagnostic assessments in endodontics (40), existing WL-related AI studies do not provide personalized, immediate, or formative feedback, which is essential for skill acquisition in a pre-clinical setting.

The precise gap addressed in this study is the absence of an AI-based educational tool that moves beyond simple image classification to deliver instant, individualized feedback on radiographic working length determination. Instant formative feedback enables students to recognize errors, self-correct, and refine their technique during the learning process; an advantage not offered by previous WL-focused AI models, which provide only categorical predictions without pedagogical guidance.

TABLE 1 Summary of AI methods used in dental radiography studies.

| Author and year | Diagnostic technique | AI method | Outcome measure |
|------------------------|---|---|--|
| Saghiri et al., 2012 | Periapical radiographs/cadaver study | ANN alone | 93% accuracy in working length determination |
| Saghiri et al., 2012 | Periapical radiographs/adaver study | ANN compared to Endodontists | 96% accuracy in anatomical position of minor apical foramen compared to 76% of endodontists |
| Qiao et al., 2020 | Circuit system designed for working length measurements | Neural network-based multi-frequency impedance method | 95% accuracy in working length determination |
| HA et al., 2023 | Radiographs/ Augmented images | Deep learning Model (YOLOv5s, YOLOv5x, and YOLOv7) | Denosing and data balancing of radiographic images have set the accuracy of the three models to be above 95% |
| Michael G et al., 2020 | Panoramic radiographs | Predictive deep learning algorithm | The rank correlation between model and cohort confidence scores for positive and negative condition cases was 0.72 and 0.34, respectively. |
| X Gao et al., 2021 | Radiographic and clinical data | BP artificial neural network model | The accuracy of this BP neural network model was 95.60% for the prediction of post-operative pain. |
| NP M et al., 2022 | Periapical radiograph (3000) | CNN to score periapical lesion on periapical radiographs using the PAI score system | True prediction PAI 1: 90.0%, PAI 2,5: 30%, PAI 3: 60%, PAI 4: 71% |
| MJ et al., 2017 | Periapical radiograph and CBCT (N = 240) | Probabilistic neural network (PNN) | PNN in CBCT images was 96.6% accurate in detecting vertical root fracture |

This study therefore aimed to develop machine learning models to provide targeted feedback on radiographic working length (WL) determination through the following steps:

- Create a new, balanced, labeled dataset of single-rooted tooth radiographs categorized into optimal, over-extended, and under-extended WL.
- Train and test 22 convolutional neural network (CNN) models on this dataset.
- Evaluate models' performance using 11 quantitative metrics.
- Select the highest-performing model—primarily based on accuracy—to develop an innovative, automated, and instant web-based WL feedback system.
- Pilot the system with pre-clinical dental students to assess its usability and educational effectiveness.

2 Material and methods

2.1 Proposed AI-driven framework

The methodology in this study outlines a detailed and systematic approach to developing a highly accurate AI-based system for working length determination using dental radiographic images in dental education. This proposed AI-driven framework incorporated both pre-trained and custom-developed Convolutional Neural Network (CNN) models, utilizing their capabilities to overcome the complexities inherent in dental radiographic images. Figure 2 presents a schematic overview of the primary components encompassed within the proposed framework. Subsequent subsections detail the various phases in this framework, providing a comprehensive roadmap

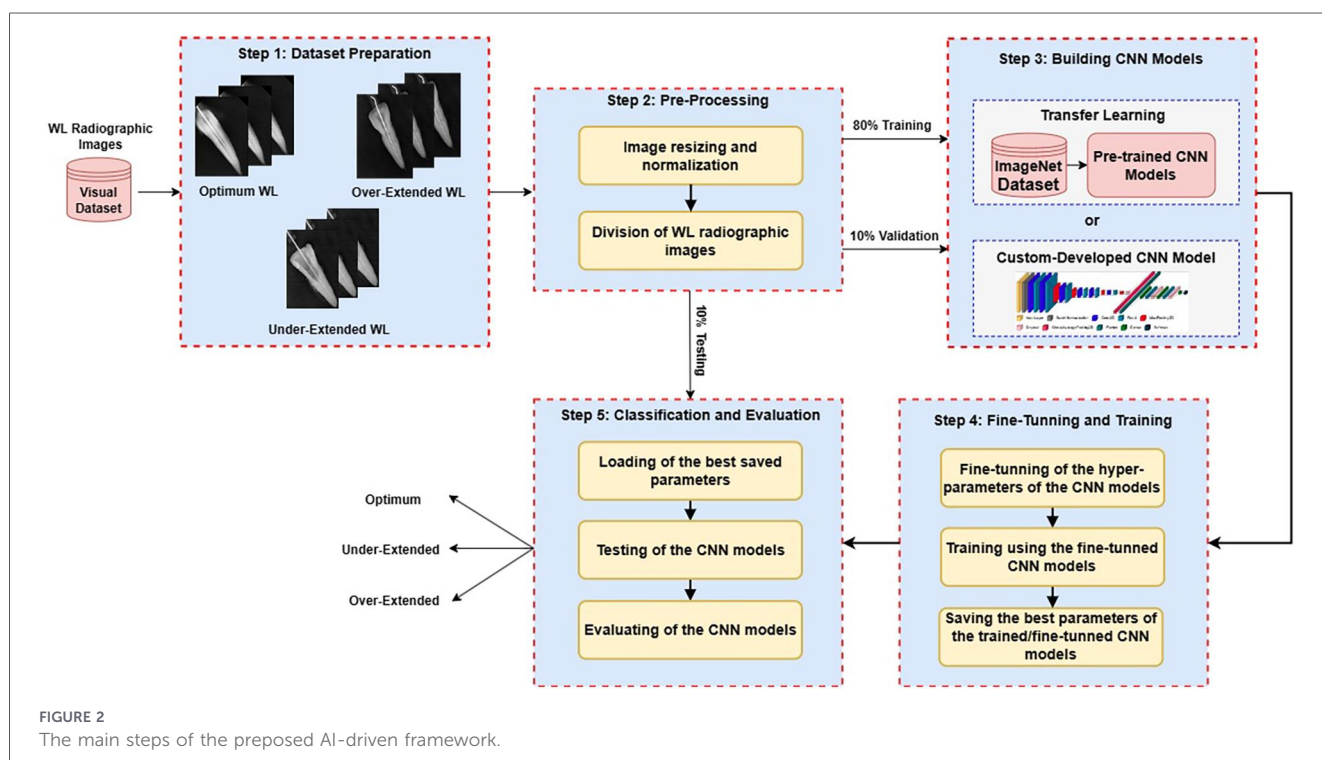
of the methodological approach adopted by the IT team in this study.

An AI-assisted feedback tool was developed and integrated within a web-based platform to provide automated evaluation and feedback on students' laboratory performance. The system was piloted with a group of 30 dental students ($\approx 10\%$ of the enrolled cohort of 300 students) who voluntarily agreed to participate towards the end of their pre-clinical year. This is justified as an appropriate pilot sample size, as 10–30 participants are considered sufficient to identify usability patterns and system acceptability.

After using the tool during their laboratory sessions, participants completed a structured feedback questionnaire to evaluate its effectiveness and usefulness. The questionnaire employed a five-point Likert scale (1 = strongly disagree to 5 = strongly agree) covering domains of usability, reliability, feedback clarity, and overall satisfaction.

2.2 Dataset preparation

The foundational step of the proposed framework is to assemble a robust dataset. This dataset initially consists of 321 high-resolution digital radiographs and has been expanded to include 3,000 digital radiographs of anterior and premolar human teeth, accessed and extracted using K files of minimum size 15, to measure the WL. These images were obtained from pre-clinical laboratory simulation exercises using extracted teeth only; no patient radiographs were used. Ethical approval for the use of extracted human teeth and radiographic imaging for research and educational purposes was obtained from the Institutional Review Board (IRB), University of Jordan (Decision No. 547/2025). All radiographs were fully de-identified prior to



analysis in accordance with IRB-approved procedures and institutional research governance requirements.

The WL dental radiographic images were collected and obtained using the advanced Ai Dental Woodpecker-V1.0.20 imaging software with acquisition parameters of (65 kVp, 7 mA, and an exposure time of 0.17 s).

Duplicate images were identified and removed to maintain data uniqueness and quality, ensuring dataset integrity, robustness, and utility. Concurrently, an extensive labeling operation was conducted, in which each image was systematically annotated to categorize the working length as optimal, under-extended, or over-extended by both a specialist endodontist and a radiologist with 4–5 years of experience. Both annotators were calibrated using 50 sample images prior to labeling the entire dataset. Inter-rater agreement was quantified using Cohen's κ , yielding a κ value of 0.77 (95% CI: 0.72–0.82), indicating substantial agreement. Disagreements were resolved through adjudication by a third senior endodontist.

The categorization protocol, developed specifically for this study, relied on precise measurements of the distance between the endodontic file tip and the radiographic root apex as recommended in the literature (4, 6, 7). The collected images were almost evenly distributed across the classifications: 1,007 radiographic images with optimum WL, 1,001 with under-extended WL, and 1,004 with over-extended WL.

Per-class counts were maintained across training, validation, and test sets using a stratified split. The final partition included 80% for training ($n = 2,400$), 10% for validation ($n = 300$), and 10% for test ($n = 300$). Splits were performed at the tooth level to prevent data leakage, ensuring no radiographs of the same tooth appeared across different subsets. Near-duplicates were removed prior to splitting. No class-specific augmentation was applied beyond standard transformations used uniformly across all categories.

Samples of WL radiographs, categorized according to the classification process, are presented in the representative images in Figure 3, with each numerical category described as follows:

- **0:** Optimum WL images, where the file tip is optimally positioned within a 0–2 mm range from the radiographic apex, ensuring that the cleaning and shaping of the root canal system is within the confines of the root structure.
- **1:** Under-extended WL images, where the file tip falls short of reaching the radiographic apex by more than 2 mm, potentially leading to inadequate disinfection.
- **2:** Over-extended WL images, where the file tip extends beyond the radiographic apex, an undesirable outcome risking over-instrumentation of the root canal system.

All WL digital radiographic images were resized to a uniform resolution of 224×224 pixels. Resizing helps reduce computational load and streamline the processing pipeline (41, 42).

Because the dataset consisted exclusively of single-rooted anterior and premolar teeth, generalizability to multi-rooted molars, anatomical variations, and different radiographic sensors may be limited. This limitation is addressed further in the discussion section. No external validation dataset from another clinic or imaging system was available for this study.

2.3 Building CNN models

This study employed a dual-pathway model development strategy which encompasses both the integration of state-of-the-art pre-trained CNN architectures [VGG16, ResNet50, VGG19, DenseNet121, DenseNet169, DenseNet201, EfficientNet series (EfficientNetB0–EfficientNetB7), InceptionResNetV2, InceptionV3, MobileNet, MobileNetV2, MobileNetV3Large, MobileNetV3Small, and Xception] (43–48), and the development of a custom CNN model designed to capture the unique features of endodontic radiographic imaging that generic and pre-trained models may not effectively capture. This aims to harness the robust capabilities of advanced neural networks while addressing the specific challenges associated with dental radiographic images.

A diverse selection of pre-trained CNN architectures was integrated, each offering distinct advantages in image recognition. VGG16 and VGG19 captured texture and details, while ResNet50 and DenseNet series addressed deep network training challenges. EfficientNet models optimized accuracy with minimal computational cost, and InceptionResNetV2 and InceptionV3 improved learning efficiency. MobileNet variants and Xception achieve a balance between speed and accuracy, making them ideal for resource-limited applications.

We began by integrating and selecting a diverse array of proven pre-trained CNN architectures, each known for its distinct advantages in various aspects of image recognition. These architectures include:

- **VGG16 and VGG19:** Renowned for their simplicity and deep architecture, which is particularly effective for capturing image texture and details.
- **ResNet50 and DenseNet Series (DenseNet121, DenseNet169, DenseNet201):** These models leverage residual and dense connections, respectively, to facilitate training deeper networks by alleviating the vanishing gradient problem.
- **EfficientNet Series (EfficientNetB0–EfficientNetB7):** Known for scaling up CNNs in a more structured manner to achieve higher accuracy without high computational costs.
- **InceptionResNetV2 and InceptionV3:** These models combine inception modules with residual connections to improve learning speed and accuracy.
- **MobileNet Series (MobileNet, MobileNetV2, MobileNetV3 Large, MobileNetV3 Small) and Xception:** Optimized for mobile devices, these models offer a good balance between speed and accuracy, making them suitable for applications where computational resources are limited.

Each of these pre-trained models is fine-tuned to adapt to our dental imaging dataset. This involves modifying the top layers of the network to focus on features specific to dental radiographs, such as tooth anatomy and the positioning of endodontic files.

To complement the broad learning capabilities of the pre-trained networks, we also developed a custom CNN architecture tailored to the nuanced requirements of endodontic radiographic analysis. The architecture of this custom model is carefully designed with several specialized layers, as demonstrated in Figure 4; Table 2:

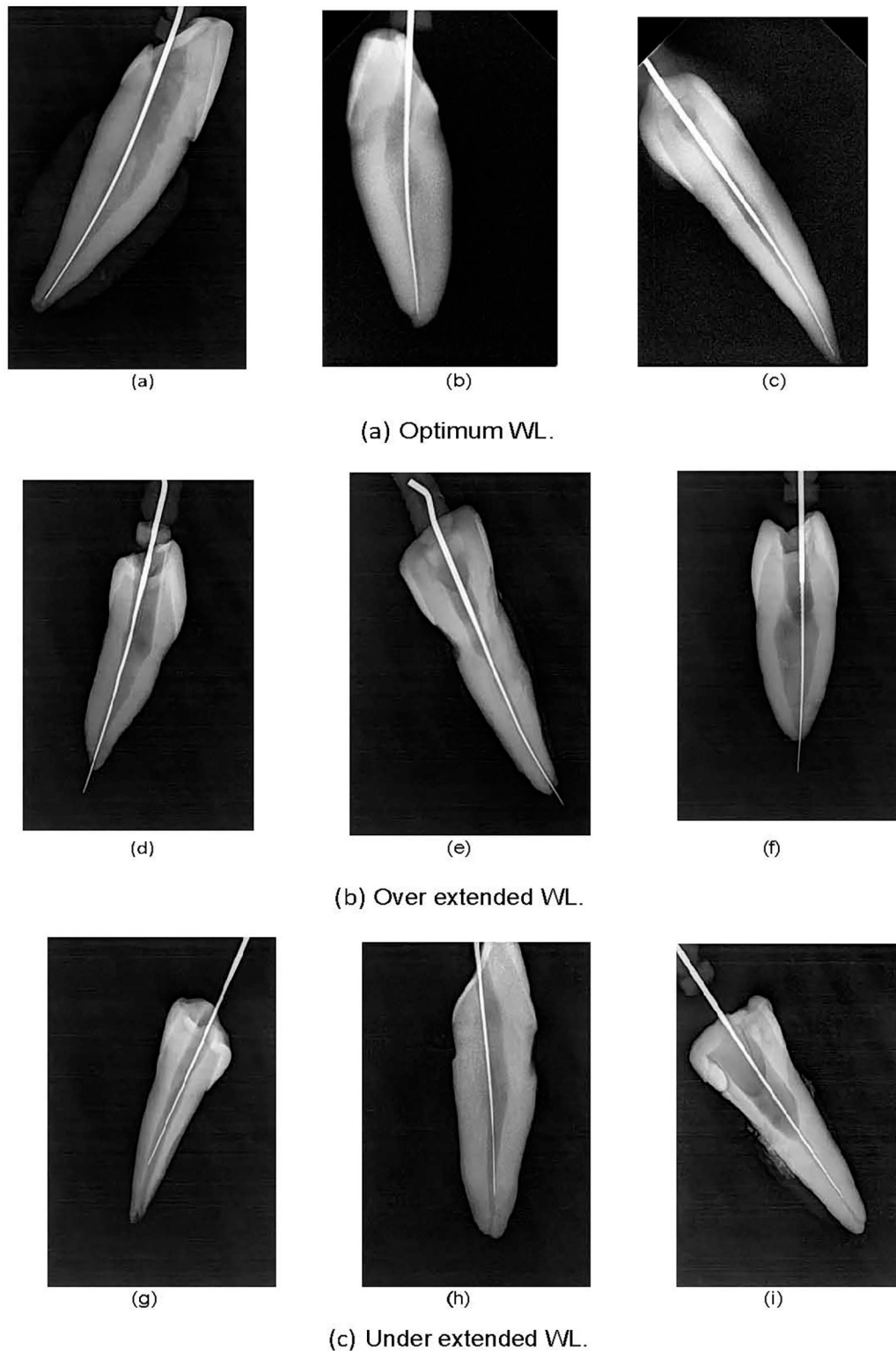


FIGURE 3

Visual samples of various examined working lengths: (a) Optimum WL, (b) over extended WL, and (c) under extended WL.

- **Convolutional Layers:** Configured to detect fine-grained details critical for accurate working length determination.
- **Depthwise Separable Convolutions:** Implemented to provide efficient model scaling and detailed feature extraction without a significant increase in computational demand.
- **Dilated Convolutions:** Used to expand the receptive field of the network, allowing it to encompass broader contextual information without losing resolution.

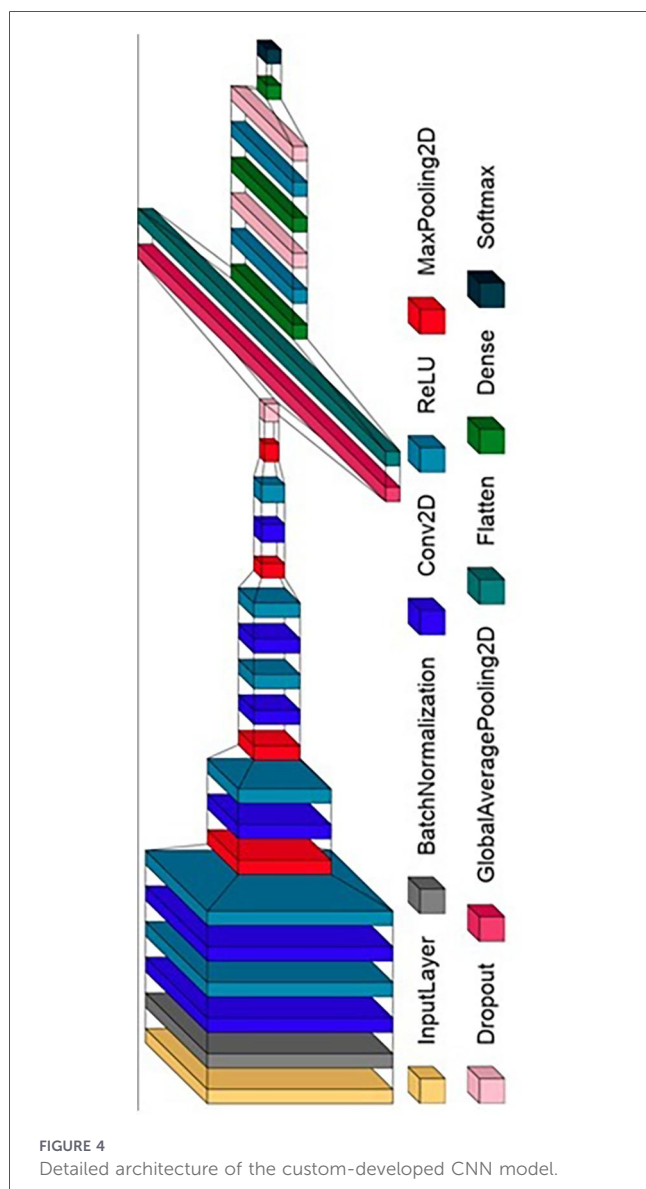


Table 2 presents the detailed layer-by-layer architecture of the custom-developed CNN model shown in Figure 4, which serves as the backbone of our WL image classification system. The initial batch normalization layer sets the stage for the network by normalizing the input distribution, which is essential for accelerating the training process.

Following this, a precisely designed sequence of 2D convolutional layers with increasing numbers of units is deployed to systematically enhance feature extraction capabilities. The strategic incorporation of max pooling layers systematically reduces the spatial dimensions of the feature maps, significantly reducing computational demands while maintaining the most salient features.

The employment of dropout layers interspersed among the high-capacity, fully connected dense layers functions as a regulatory mechanism to combat overfitting, ensuring that the model generalizes well to new, unseen data. Activation functions such as ReLU introduce nonlinearity, enabling the model to learn complex patterns. Furthermore, batch normalization layers are incorporated to enhance training

stability, and dropout is strategically applied to prevent overfitting.

The final output layer, with its SoftMax activation, is calibrated to output three classes, translating the learned patterns into clinically relevant classifications of working length (optimum, over-extended, or under-extended). The proposed CNN architecture has 1.5 million parameters, balancing model complexity and interpretability, a balance paramount in medical image analysis.

By deploying this dual approach, combining the generalized learning capabilities of pre-trained models with the focused precision of a custom-developed architecture, we aim to create a robust system capable of high accuracy and reliability in classifying dental radiographic images.

This methodology not only leverages the strengths of both architectural paradigms but also ensures the models are finely tuned to meet the specific challenges posed by endodontic imaging tasks.

2.4 Fine-tuning and training

The fine-tuning and training of the selected and developed CNN models are critical stages in our study, during which we refine the models to ensure optimal performance on dental radiographic images. This phase is thoroughly designed to fine-tune both pre-trained and custom-developed models, enhancing their ability to accurately identify and classify the varied anatomical features and pathologies evident in endodontic imaging.

Fine-tuning is approached by unfreezing the top layers of the pre-trained models, which are the most likely to capture high-level features specific to our dental dataset. The parameters of these layers are then carefully adjusted through continued training on our dataset (49). We modify:

- **Learning Rates:** Initially set lower than usual to prevent the loss of previously learned features. The learning rate is progressively increased using a cyclical policy, which helps escape local minima and explore a broader region of the solution space.
- **Layer Configurations:** Depending on the specific model architecture, certain layers may be added or removed to better capture the nuances of dental images. For instance, additional convolutional layers may be added to deepen the model's ability to process fine details, or fully connected layers may be adjusted to refine the output towards the specific classes of interest.

Once the models are fine-tuned, they undergo a rigorous training regime using the following advanced techniques to enhance their generalizability and robustness:

- **Dropout:** Applied randomly to neurons in the dense layers during training to prevent overfitting by reducing interdependencies among neurons. This method effectively improves the model's ability to generalize from the training data to unseen data.

TABLE 2 Layer-by-layer architecture of the custom-developed CNN model.

| Type of the layer | Output configuration | Parameter characteristics |
|--|---|---|
| Batch Normalization (batch normalization) | 224 × 224 spatial resolution, 3 channels | Count: 12, Normalization parameters |
| 2D Convolution (conv2d) | 224 × 224 spatial resolution, 8 units | Count: 224, Feature detectors |
| 2D Convolution (conv2d 1) | 224 × 224 spatial resolution, 16 units | Count: 1,168, Feature detectors |
| Max Pooling (max pooling2d) | Halved resolution: 112 × 112, 16 units | None, Dimensionality reduction |
| 2D Convolution (conv2d 2) | Halved resolution: 112 × 112, 32 units | Count: 4,640, Feature detectors |
| Max Pooling (max pooling2d 1) | Quartered resolution: 56 × 56, 32 units | None, Dimensionality reduction |
| 2D Convolution (conv2d 3) | Quartered resolution: 56 × 56, 64 units | Count: 18,496, Feature detectors |
| 2D Convolution (conv2d 4) | Quartered resolution: 56 × 56, 64 units | Count: 36,928, Feature detectors |
| Max Pooling (max pooling2d 2) | One-eighth resolution: 28 × 28, 64 units | None, Dimensionality reduction |
| 2D Convolution (conv2d 5) | One-eighth resolution: 28 × 28, 256 units | Count: 147,712, Advanced feature extraction |
| Max Pooling (max pooling2d 3) | One-sixteenth resolution: 14 × 14, 256 units | None, Dimensionality reduction |
| Dropout (dropout) | Preserved resolution: 14 × 14, 256 units | None, Overfitting mitigation |
| Global Average Pooling (global average pooling2d) | Compressed feature representation: 256 units | None, Feature summarization |
| Flatten (flatten) | Vectorization of features: 256 units | None, Preparing for dense layers |
| Fully Connected (dense) | High-capacity feature processing: 1,024 units | Count: 263,168, Learning dense features |
| Dropout (dropout 1) | Feature selection: 1,024 units | None, Overfitting mitigation |
| Fully Connected (dense 1) | High-capacity feature processing: 1,024 units | Count: 1,049,600, Learning dense features |
| Dropout (dropout 2) | Feature selection: 1,024 units | None, Overfitting mitigation |
| Output Layer (dense 2) | Final decision layer: 3 units (class probabilities) | Count: 10,250, Classification decisions |
| Total parameters: 1,532,198, Comprising Trainable: 1,532,192, Non-trainable: 6 | | |

- **Batch Normalization:** Implemented after every convolutional layer to normalize the activations and accelerate the training process. This technique also helps stabilize the neural network by normalizing the input layer by adjusting the mean and variance.

Both dropout and batch normalization are pivotal for managing internal covariate shift, thereby speeding up training and improving the model's performance during inference on new, unseen images.

Training is conducted in batches, with batch size accurately chosen based on the computational capabilities and the specific architecture being trained. Gradient descent algorithms, specifically Adam or Stochastic Gradient Descent (SGD) with momentum, are utilized to update the weights. These optimizers are selected for their efficiency in handling sparse gradients and robustness in diverse training landscape conditions.

Throughout the training process, the model's performance is continuously monitored using a validation set. This not only ensures the model is not overfit with the training data but also provides insights into how it might perform in a real-world clinical setting. Based on the validation results, the training parameters are iteratively refined to further enhance accuracy and reliability.

Table 3 presents the fine-tuned configuration parameters used to train the convolutional neural networks described in this work. It highlights the software environment and tools, such as Python with TensorFlow and Keras, and emphasizes the advanced computational framework employed. The data is

TABLE 3 Detailed fine-tuned configuration parameters for CNN model training.

| Parameter | Specification |
|----------------------------------|---------------------------|
| Development language | Python |
| Neural network libraries | TensorFlow, Keras |
| Data split (Train/Validate/Test) | 80%/10%/10% |
| Learning rate | 0.0001 |
| Optimizer | Adam |
| Regularization technique | L2 regularization |
| L2 regularization rate | 0.001 |
| Training epochs | 128 |
| Batch size | 64 |
| Loss computation | Categorical cross-entropy |

divided into training, validation, and test sets at 80%, 10%, and 10%, respectively, facilitating robust model training and evaluation. A learning rate of 0.0001, combined with the Adam optimization algorithm, ensures efficient convergence during training. Regularization is handled via an L2 penalty with a decay rate of 0.001 to prevent overfitting, thereby improving the models' generalization. The networks are trained for 128 epochs with a minimum batch size of 64, using a categorical cross-entropy loss function, which is optimal for the multi-class classification problems inherent to this study.

TABLE 4 The detailed evaluation metrics and their definitions.

| Evaluation metric | Definition |
|---|--|
| Accuracy | Measures the overall correctness of the model across all classes, giving us a straightforward indication of performance |
| Precision (Positive Predictive Value) | Indicates the accuracy of positive predictions for each class, crucial for applications where the cost of a false positive is high. |
| Recall (Sensitivity or True Positive Rate): | Measures the model's ability to detect all positive instances, which is essential in clinical settings where missing a positive case can have profound implications. |
| F1-score | Harmonizes precision and recall into a single metric, balancing the trade-offs between them and providing a more comprehensive overview of model performance. |
| Testing time | Indicates the efficiency of the model, ensuring the model can be integrated smoothly to provide output without significant delays. |
| Specificity (True negative rate) | Measures the proportion of actual negatives that are correctly identified as such, necessary for confirming the absence of a condition. |
| Negative predictive value (NVP) | Indicates the likelihood that a negative test truly means the absence of a condition, which is valuable in ensuring the reassurance of negative screening results. |
| False positive rate (FPR) | Indicates the probability of falsely classifying a negative case as positive, critical in areas where false alarms are costly. |
| False negative rate (FNR) | Represents the probability of falsely classifying a positive case as negative, which is crucial in settings where missing a condition could be detrimental. |
| False discovery rate (FDR) | Measures the proportion of false positives in all positive predictions, which is essential for understanding the reliability of positive test results. |
| False omission rate (FOR) | Measures the proportion of false negatives in all negative predictions, relevant in scenarios where negatives carry significant implications. |
| Misclassification rate (MR) | Measures the overall rate at which the model incorrectly classifies instances, giving an overall error rate of the system. |

This comprehensive fine-tuning and training process is essential for optimizing the performance of our CNN models, ensuring they are not only theoretically sound but also practically viable and effective in a clinical environment.

2.5 CNN models' evaluation and classification process

The evaluation process for CNN models assesses their accuracy and reliability to determine their effectiveness and readiness for deployment in clinical settings. The 22 CNN models were comprehensively evaluated using 11 different metrics that are listed in Table 4. The model with the best performance was then plugged into a web-based tool to offer a system with high accuracy, reliability, and practical usability in clinical applications.

3 Results and analysis

A comprehensive analysis of the outcomes derived from the detailed evaluation of the CNN models developed to classify working lengths in dental radiographs. Key performance metrics, including accuracy, F1-score, precision, and recall, were computed for all examined models, along with testing time to assess operational efficiency. An in-depth detection assessment was conducted on the most accurate custom-developed CNN model, encompassing all relevant evaluation parameters across the three designated WL categories. In terms of precision, the DCNN model achieved the highest score of 97.07%, closely followed by EfficientNetB5 and VGG16, also known for their

deep learning efficiency in image recognition tasks. The recall metric, critical for ensuring correct classification, was also highest for the DCNN model at 97.03%. The DCNN model excelled in both accuracy and operational efficiency, achieving a test time of only 0.548546 s. EfficientNetB5, VGG16, EfficientNetB1, VGG19, and EfficientNetB2 also achieved commendable accuracy, with scores above 95%, underscoring the efficacy of advanced CNN architectures in handling complex image data. Table 5 presents the detailed detection assessment parameters for each examined AI model.

The custom-developed deep CNN (DCNN) (Scratch) model achieved superior performance, with an accuracy of 97.02% and an F1-score of 97.03%, indicating highly reliable classification. Table 6 outlines the detection assessment parameters for the DCNN model for the optimum, over-extended, and under-extended WL categories.

Additionally, to conduct a deep analysis, we evaluated the performance of the top six CNN models based on their accuracy and loss curves, and the precision of their classification capabilities as depicted in their confusion matrices. These analyses provide insights into the models' abilities to learn and generalize from the training data, as well as their precision in classifying working lengths.

The accuracy and loss curves, illustrated in Figure 5 [(a)–(f)], represent key indicators of model performance over training epochs for the top-performing models such as DCNN, EfficientNetB5, VGG16, EfficientNetB1, VGG19, and EfficientNetB2.

In Figure 5(a), a slight increase in the validation loss is observed after approximately epoch 40. This corresponds to the inflection point where the DCNN achieves its optimal generalization. Beyond this stage, further minimization of the

TABLE 5 Detection assessment parameters of the examined AI models.

| AI model | Accuracy | F1-score | Precision | Recall | Testing time (sec) |
|-------------------|-----------------|-----------------|-----------------|-----------------|--------------------|
| DCNN (Scratch) | 0.970297 | 0.970390 | 0.970702 | 0.970297 | 0.548546 |
| EfficientNetB5 | 0.960396 | 0.960613 | 0.961948 | 0.960396 | 6.703499 |
| VGG16 | 0.957096 | 0.957219 | 0.957729 | 0.957096 | 7.026967 |
| EfficientNetB1 | 0.957096 | 0.957182 | 0.957879 | 0.957095 | 2.888323 |
| VGG19 | 0.950495 | 0.950694 | 0.951500 | 0.950495 | 9.019332 |
| EfficientNetB2 | 0.953795 | 0.953937 | 0.954361 | 0.953795 | 3.007454 |
| EfficientNetB7 | 0.947195 | 0.947434 | 0.947956 | 0.947195 | 12.19628 |
| MobileNetV3Large | 0.947195 | 0.947207 | 0.948072 | 0.947195 | 1.444528 |
| EfficientNetB3 | 0.940594 | 0.940899 | 0.941771 | 0.940594 | 3.825811 |
| EfficientNetB6 | 0.943894 | 0.943810 | 0.944299 | 0.943894 | 9.057463 |
| ResNet50 | 0.930693 | 0.931268 | 0.933418 | 0.930693 | 4.158439 |
| MobileNet | 0.927393 | 0.927840 | 0.928913 | 0.927393 | 1.177610 |
| EfficientNetB0 | 0.927393 | 0.927491 | 0.928669 | 0.927393 | 2.072861 |
| DenseNet121 | 0.927393 | 0.927351 | 0.927517 | 0.927392 | 4.216988 |
| EfficientNetB4 | 0.927393 | 0.927691 | 0.928396 | 0.927393 | 4.960827 |
| MobileNetV3Small | 0.920792 | 0.920765 | 0.921505 | 0.920792 | 0.878270 |
| DenseNet169 | 0.920792 | 0.916644 | 0.922791 | 0.920792 | 5.148977 |
| DenseNet201 | 0.920792 | 0.920544 | 0.922299 | 0.920792 | 6.699370 |
| Xception | 0.910891 | 0.910857 | 0.911724 | 0.910891 | 3.529581 |
| InceptionV3 | 0.891089 | 0.891423 | 0.892822 | 0.891089 | 2.156583 |
| MobileNetV2 | 0.864686 | 0.864373 | 0.868220 | 0.864686 | 1.380209 |
| InceptionResNetV2 | 0.617161 | 0.614039 | 0.614189 | 0.617161 | 5.616386 |

TABLE 6 Detailed detection assessment parameters of the most accurate DCNN model across the three examined working length samples.

| Metric | Optimum WL | Over-extended WL | Under-extended WL |
|----------|------------|------------------|-------------------|
| Acc. | 0.970297 | 0.990099 | 0.980198 |
| TNR | 0.970297 | 0.995050 | 0.990099 |
| NPV | 0.984925 | 0.990147 | 0.980392 |
| FPR | 0.029703 | 0.004950 | 0.009901 |
| FNR | 0.029702 | 0.019801 | 0.039604 |
| PPV | 0.942307 | 0.990000 | 0.979798 |
| FDR | 0.057692 | 0.010000 | 0.020202 |
| TPR | 0.970297 | 0.980198 | 0.960396 |
| FOR | 0.015075 | 0.009852 | 0.019608 |
| F1-score | 0.956097 | 0.985074 | 0.970000 |
| MR | 0.029703 | 0.009900 | 0.019802 |

training objective can lead to mild overfitting, as the learned feature representations become increasingly tailored to the training data. Moreover, because categorical cross-entropy is highly sensitive to prediction confidence, small variations in probability estimates for challenging validation instances can cause the loss value to rise without a corresponding decrease in classification accuracy. To ensure optimal performance, the

system used a model checkpointing approach, saving and loading the best parameters identified before this trend emerged, thereby maintaining the stability and reliability reported in the final evaluation metrics.

Notably, the DCNN model exhibits rapid stabilization in accuracy and a corresponding decrease in loss, achieving both high training accuracy and low validation loss early in training. This suggests an effective learning process with minimal overfitting. EfficientNetB5 and VGG16 also show exemplary convergence behavior, indicating robustness in their learning mechanisms.

Figure 6 [(a)–(f)] displays the confusion matrices for the same subset of top-performing CNN models, offering a breakdown of their predictive performance across the three classified working lengths: optimum, over-extended, and under-extended. The matrices show high true-positive rates across all six CNN models, with the DCNN achieving near-perfect classification across all categories. The EfficientNetB5 model shows a slight misclassification between the optimal and over-extended categories, suggesting opportunities to improve at distinguishing closely related classes.

4 Web-based smart WL feedback system: case study

The best AI model was integrated into a web-based service that allows students to submit their work (WL

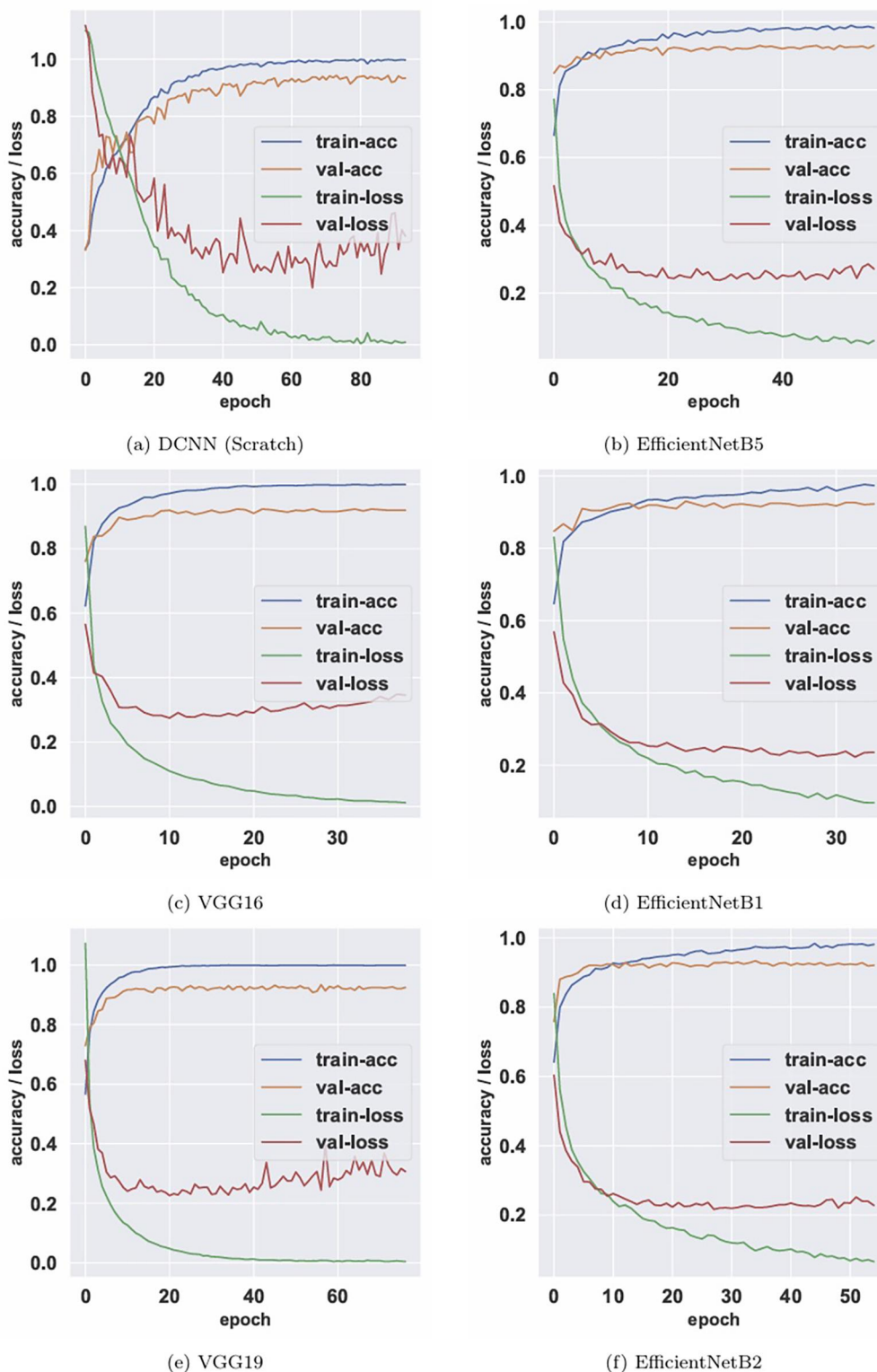


FIGURE 5 Accuracy and loss curves for the top six most accurate AI models.

radiograph) and receive prompt, constructive feedback without delay. For instance, consider the three different inputs from three students, as shown in Figure 7. The feedback coming from our automated smart system was as follows:

- For (code 0), Figure 7A, the system classifies it as an optimal position since the file tip is within 2 millimeters (mm) of the radiographic apex. The student is then instructed to proceed to the next step, which involves chemo-mechanical root canal preparation.

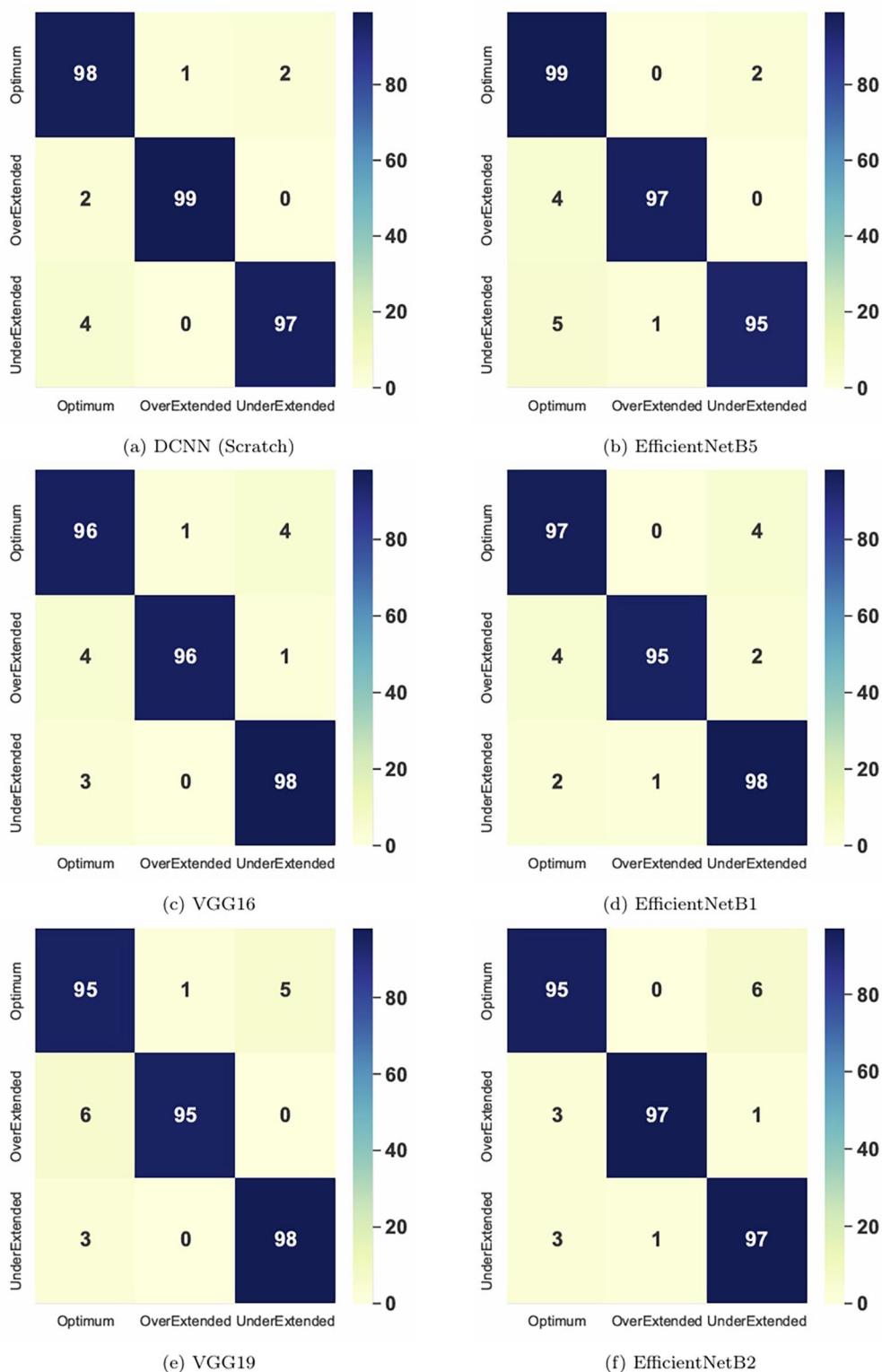


FIGURE 6 Confusion matrices for the top six most accurate AI models.

- For (code 1), Figure 7B, the system classifies it as under-extended since the file tip falls short of the radiographic apex by more than 2 mm. The student is then instructed to adjust the file length by increasing it, re-inserting it, and taking a new radiograph.
- For (code 2), Figure 7C, the system classifies it as over-extended since the file tip extends beyond the radiographic apex by more than 2 mm. The student is then instructed to adjust the file length by reducing it, re-inserting it, and taking a new radiograph.



TABLE 7 Summary of student responses to the AI-assisted feedback tool (n = 30).

| Qualitative study results analysis | | |
|--|-----------------|----------------|
| Statement | Mean \pm SD | Interpretation |
| The tool was easy to set up and use | 4.71 \pm 0.46 | Strongly agree |
| Instructions for using the tool were clear and sufficient | 4.74 \pm 0.44 | Strongly agree |
| The tool functioned reliably without technical difficulties | 4.68 \pm 0.47 | Strongly agree |
| Using the tool allowed me to complete the laboratory tasks efficiently | 4.55 \pm 0.51 | Agree |

Thirty dental students completed the survey evaluating the AI-assisted laboratory feedback tool. These 30 students, representing approximately 10% of the total enrolled pre-clinical cohort, were selected for piloting purposes.

The sampling frame included all pre-clinical students at the end of their laboratory training period. Inclusion criteria were A) completion of the pre-clinical endodontic laboratory course, and B) willingness to test the web-based tool. Exclusion criteria included: A) incomplete laboratory attendance, or B) prior participation in tool development. No compensation or incentives were offered; participation was voluntary.

The overall mean rating was 4.67/5, indicating strong satisfaction. The highest-rated item was 'Instructions for using the tool were clear and sufficient' (mean = 4.74). The remaining rates for the other items questioned in the survey are listed in Table 7. However, because Likert data are ordinal, the median score across all items was 5.0, with IQR values ranging from 4 to 5 to 5–5.

Internal consistency of the questionnaire was assessed using Cronbach's alpha, which demonstrated excellent reliability ($\alpha = 0.96$).

The questionnaire consisted of 10 items covering usability, clarity of instructions, reliability, feedback quality, and perceived learning enhancement. Items were scored using a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). Although this was a custom instrument tailored to the WL determination context, a standardized metric such as SUS or UMUX-Lite will be considered in future work for broader comparability.

Sixteen female students and fourteen males participated in this survey. The web-based tool is hosted on a web server, and its link was shared with the relevant parties. Students consistently agreed that the tool was easy to use, provided clear and timely feedback, and supported efficient task completion. Overall, the tool was perceived as reliable, user-friendly, and beneficial for enhancing learning in endodontic laboratory training.

5 Discussions

In Endodontics, AI is increasingly integrated into the technical assessment of root canal treatment through radiographic examination (19, 26, 36, 50, 51). Stages such as WL determination are critical steps in root canal treatment and

directly affect its success (6, 7, 32). Providing personalized feedback to each case during endodontic training can be challenging, but it is essential for enhancing effective learning, especially in large classrooms (33, 52, 53).

This experimental study aimed to significantly enhance the precision and reliability of endodontic pre-clinical training by integrating AI Machine learning models with dental radiography. This proposed model incorporated both pre-trained and custom-developed 22 CNN models. The matrices show high true-positive rates across all six CNN models, with the DCNN achieving near-perfect classification across all categories.

The efficiency of clinical training in dental education, specifically within the context of endodontic training protocols, is fundamentally contingent upon accurately determining working length (WL) 4,6,7. Leveraging advanced CNN capabilities, this study introduces an AI-driven framework that enhances the precision of these measurements.

This precise labeling protocol was critical for training the CNN models, as it encapsulated various scenarios a dental practitioner may encounter, thereby enhancing their accuracy and reliability. A well-balanced and refined dataset of 3,000 samples was created, with care taken to maintain consistent image quality and resolution, ensuring each radiograph was suitable for detailed analysis and machine learning applications. The categorization of radiographic WL as optimal, under-extended, and over-extended WL was considered based on the recommendations for radiographic WL measurements in the literature (6–8). In addition, the distribution of images was balanced across the three categories to avoid dataset bias and enhance the validity of the AI-based learning process.

The DCNN model excelled in both accuracy and operational efficiency, achieving a test time of only 0.5 s. This rapid testing capability is advantageous for real-time applications, enabling the model to be integrated smoothly and providing output without significant delays. This high recall rate is essential in clinical settings where missing or delaying diagnostic information could lead to poor procedural outcomes.

This high accuracy and balanced F1 score suggest that the model effectively manages the trade-off between precision and recall, which is crucial for medical applications where false positives and false negatives carry significant consequences.

The accuracy of the DCNN AI-driven model developed in this study (reaching up to 99%) exceeded the accuracy of other AI-driven models exploring WL determination (ranging from 93% to 96%) (29, 30, 54). This could be justified by the integration of both empirical and advanced AI techniques in this study, in addition to the robust and large dataset used to train this model. The use of a case study for a trial of this model was proven to provide effective feedback to the three main categories of WL radiographs. This indicates that automated, accurate, and quick feedback responses will be provided to students, assisting and supporting instructors in offering constructive feedback during this critical step (33, 55, 56). The AI-based feedback model developed through this work could be employed to deliver instant, precise, and constructive feedback to students. This custom model includes specialized layers and training strategies designed to enhance feature extraction directly pertinent to the nuances of endodontic imaging.

This study represents a laboratory-based exploration of AI and machine learning development, making it one of the first initiatives in this direction. The use of a large, well-organized, and quality-assured dataset of WL radiographic images has contributed to the development of a robust model. The high satisfaction scores indicate that the AI-assisted feedback tool was well accepted by students and effectively supported their learning in pre-clinical endodontic sessions. The tool's clarity, ease of use, and instant feedback appear to enhance both efficiency and confidence during laboratory tasks. Such positive perceptions support the continued integration of AI-driven systems into pre-clinical teaching to promote active and self-directed learning.

The limitations of this study must be acknowledged. Since all radiographs were obtained from a single institution using a single imaging software and device, this can limit generalizability across clinics using different radiographic sensors with varying specs.

Although the model performed well diagnostically, this study did not test whether AI-supported feedback improves learning outcomes. Future controlled research, such as a randomized comparison with standard instruction, is required to determine educational effectiveness. This is an alpha version of the tool, and we are improving it.

6 Conclusions

This study presents the development of an AI-powered automated feedback system designed to enhance pre-clinical endodontic training by assisting students in determining radiographic working length (WL). A newly labeled dataset of 3,000 radiographs was used to train 22 convolutional neural network (CNN) models, achieving up to 99% accuracy, with the best-performing model selected for implementation. The system provides instant, constructive feedback to support student learning and technical improvement. Future work will focus on expanding the dataset, testing additional AI architectures, and integrating other stages of root canal treatment into a unified intelligent training platform for dental education.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

- Carrotte PV. An introduction to endodontics. *BDJ Team*. (2021) 8(4):31–4. doi: 10.1038/s41407-021-0582-3
- Kakehashi S, Stanley H, Fitzgerald R. The effects of surgical exposures of dental pulps in germ-free and conventional laboratory rats. *Oral Surg Oral Med Oral Pathol*. (1965) 20(3):340–9. doi: 10.1016/0030-4220(65)90166-0
- Estrela C, Holland R, Estrela C, Alencar AHG, Sousa-Neto MD, Pécora JD. Characterization of successful root canal treatment. *Braz Dent J*. (2014) 25:3–11. doi: 10.1590/0103-6440201302356
- Bhatt A, Gupta V, Rajkumar B, Arora R. Working length determination-the soul of root canal therapy a review. *Int J Dent Helt Sci*. (2015) 2(1):105–15.
- Schilder H. Cleaning and shaping the root canal. *Dent Clin North Am*. (1974) 18(2):269–96. doi: 10.1016/S0011-8532(22)00677-2
- Sjögren U, Hägglund B, Sundqvist G, Wing K. Factors affecting the long-term results of endodontic treatment. *J Endod*. (1990) 16(10):498–504. doi: 10.1016/S0099-2399(07)80180-4
- Ricucci L. Apical limit of root canal instrumentation and obturation, part 2. A histological study. *Int Endod J*. (1998) 31(6):394–409. doi: 10.1046/j.1365-2591.1998.00183.x
- Gordon M, Chandler N. Electronic apex locators. *Int Endod J*. (2004) 37(7):425–37. doi: 10.1111/j.1365-2591.2004.00835.x

Author contributions

SA: Writing – original draft, Resources, Conceptualization, Writing – review & editing, Data curation, Supervision, Project administration. IA: Writing – original draft, Methodology, Conceptualization, Visualization, Validation, Writing – review & editing, Supervision, Formal analysis, Project administration, Resources. WE-S: Formal analysis, Software, Validation, Visualization, Writing – original draft. MA-A: Conceptualization, Methodology, Validation, Visualization, Writing – original draft. AA: Data curation, Writing – original draft. RA: Data curation, Formal analysis, Writing – original draft.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

9. Ng YL, Mann V, Gulabivala K. Tooth survival following non-surgical root canal treatment: a systematic review of the literature. *Int Endod J.* (2010) 43(3):171–89. doi: 10.1111/j.1365-2591.2009.01671.x
10. Zadik Y, Sandler V, Bechor R, Salehrabi R. Analysis of factors related to extraction of endodontically treated teeth. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontology.* (2008) 106(5):e31–5. doi: 10.1016/j.tripleo.2008.06.017
11. Stabholz A, Rotstein I, Torabinejad M. Effect of preflaring on tactile detection of the apical constriction. *J Endod.* (1995) 21(2):92–4. doi: 10.1016/S0099-2399(06)81103-9
12. Chandler N, Bloxham G. Effect of gloves on tactile discrimination using an endodontic model. *Int Endod J.* (1990) 23(2):97–9. doi: 10.1111/j.1365-2591.1990.tb00846.x
13. Jenkins JA, Walker WA III, Schindler WG, Flores CM. An *in vitro* evaluation of the accuracy of the root ZX in the presence of various irrigants. *J Endod.* (2001) 27(3):209–11. doi: 10.1097/00004770-200103000-00018
14. Trope M, Rabie G, Tronstad L. Accuracy of an electronic apex locator under controlled clinical conditions. *Dent Traumatol.* (1985) 1(4):142–5. doi: 10.1111/j.1600-9657.1985.tb00579.x
15. Endodontology ESo. Quality guidelines for endodontic treatment: consensus report of the European society of endodontology. *Int Endod J.* (2006) 39(12):921–30. doi: 10.1111/j.1365-2591.2006.01180.x
16. Orafi I, Rushton V. The use of radiography and the apex locator in endodontic treatment within the UK: a comparison between endodontic specialists and general dental practitioners. *Int Endod J.* (2013) 46(4):355–64. doi: 10.1111/j.1365-2591.2012.02127.x
17. Ahmed ZH, Almuharib AM, Abdulkarim AA, Alhassoon AH, Alanazi AF, Alhaqbani MA, et al. Artificial intelligence and its application in endodontics: a review. *J Contemp Dent Pract.* (2024) 24(11):912–7. doi: 10.5005/jp-journals-10024-3593
18. Bonny T, Al Nassan W, Obaideen K, Al Mallahi MN, Mohammad Y, El-Damanhoury HM. Contemporary role and applications of artificial intelligence in dentistry. *FI000Res.* (2023) 12:1–27. doi: 10.12688/fi000research.140204.1
19. Endres MG, Hillen F, Salloum M, Sedaghat AR, Niehues SM, Quatela O, et al. Development of a deep learning algorithm for periapical disease detection in dental radiographs. *Diagnostics.* (2020) 10(6):430. doi: 10.3390/diagnostics10060430
20. Marwaha J. Artificial intelligence in conservative dentistry and endodontics: a game-changer. *J Conserv Dentistry Endodontics.* (2023) 26(5):514–8. doi: 10.4103/JCDE.JCDE_7_23
21. Prados-Privado M, García Villalón J, Martínez-Martínez CH, Ivorra C, Prados-Frutos JC. Dental caries diagnosis and detection using neural networks: a systematic review. *J Clin Med.* (2020) 9(11):3579. doi: 10.3390/jcm9113579
22. Lee J, Seo H, Choi YJ, Lee C, Kim S, Lee YS, et al. An endodontic forecasting model based on the analysis of preoperative dental radiographs: a pilot study on an endodontic predictive deep neural network. *J Endod.* (2023) 49(6):710–9. doi: 10.1016/j.joen.2023.03.015
23. Orhan K, Bayraktar I, Ezhov M, Kravtsov A, Özyürek T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. *Int Endod J.* (2020) 53(5):680–9. doi: 10.1111/iej.13265
24. Johari M, Esmaili F, Andalib A, Garjani S, Saberhari H. Detection of vertical root fractures in intact and endodontically treated premolar teeth by designing a probabilistic neural network: an *ex vivo* study. *Dentomaxillofacial Radiology.* (2017) 46(2):20160107. doi: 10.1259/dmfr.20160107
25. Nardi C, Calistri L, Pradella S, Desideri I, Lorini C, Colagrande S. Accuracy of orthopantomography for apical periodontitis without endodontic treatment. *J Endod.* (2017) 43(10):1640–6. doi: 10.1016/j.joen.2017.06.020
26. Hasan HA, Saad FH, Ahmed S, Mohammed N, Farook TH, Dudley J. Experimental validation of computer-vision methods for the successful detection of endodontic treatment obturation and progression from noisy radiographs. *Oral Radiol.* (2023) 39(4):683–98. doi: 10.1007/s11282-023-00685-8
27. Khanagar SB, Al-Ehaideb A, Vishwanathiah S, Maganur PC, Patil S, Naik S, et al. Scope and performance of artificial intelligence technology in orthodontic diagnosis, treatment planning, and clinical decision-making: a systematic review. *J Dent Sci.* (2021) 16(1):482–92. doi: 10.1016/j.jds.2020.05.022
28. Karobari MI, Adil AH, Basheer SN, Murugesan S, Savadamoorthi KS, Mustafa M, et al. Evaluation of the diagnostic and prognostic accuracy of artificial intelligence in endodontic dentistry: a comprehensive review of literature. *Comput Math Methods Med.* (2023) 2023:1–9. doi: 10.1155/2023/7049360
29. Saghir MA, Asgar K, Boukani KK, Lotfi M, Aghili H, Delvarani A, et al. A new approach for locating the minor apical foramen using an artificial neural network. *Int Endod J.* (2012) 45(3):257–65. doi: 10.1111/j.1365-2591.2011.01970.x
30. Qiao X, Zhang Z, Chen X. Multifrequency impedance method based on neural network for root canal length measurement. *Appl Sci.* (2020) 10(21):7430. doi: 10.3390/app10217430
31. Seijo MO, Ferreira EF, Ribeiro Sobrinho AP, Paiva SM, Martins RC. Learning experience in endodontics: brazilian students' perceptions. *J Dent Educ.* (2013) 77(5):648–55. doi: 10.1002/j.0022-0337.2013.77.5.tb05515.x
32. Baaij A, Kruse C, Whitworth J, Jarad F. European Society of endodontology undergraduate curriculum guidelines for endodontology. *Int Endod J.* (2024) 57(8):982–95. doi: 10.1111/iej.14064
33. Davey J, Bryant S, Dummer P. The confidence of undergraduate dental students when performing root canal treatment and their perception of the quality of endodontic education. *Eur J Dent Educ.* (2015) 19(4):229–34. doi: 10.1111/eje.12130
34. Hattar S, AlHadidi A, Altarawneh S, Hamdan AA, Shaini FJ, Wahab FK. Dental students' experience and perceived confidence level in different restorative procedures. *Eur J Dent Educ.* (2021) 25(1):207–14. doi: 10.1111/eje.12592
35. Baaij A, Özok A. Method of teaching undergraduate students to perform root canal treatment: it's influence on the quality of root fillings. *Eur J Dent Educ.* (2018) 22(2):e221–7. doi: 10.1111/eje.12275
36. Moidu NP, Sharma S, Chawla A, Kumar V, Logani A. Deep learning for categorization of endodontic lesion based on radiographic periapical index scoring system. *Clin Oral Investig.* (2022) 26(1):651–8. doi: 10.1007/s00784-021-04043-y
37. Yüzbaşıoğlu E. Attitudes and perceptions of dental students towards artificial intelligence. *J Dent Educ.* (2021) 85(1):60–8. doi: 10.1002/jdd.12385
38. Buyuk C. Should artificial intelligence integrate with dental education? An assessment through the dentomaxillofacial radiology perspective. *J Res Dentistry.* (2021) 9(1):6–13. doi: 10.19177/jrd.v9e120216-13
39. Nassar HM, Tekian A. Computer simulation and virtual reality in undergraduate operative and restorative dental education: a critical review. *J Dent Educ.* (2020) 84(7):812–29. doi: 10.1002/jdd.12138
40. Prinz M, Bürklein S, Schäfer E, Donnermeyer D. An AI-based e-learning tool to improve endodontic diagnostics in undergraduate students. *J Dent Educ.* (2024) 88:1935–37. doi: 10.1002/jdd.13479
41. El-Shafai W, Almomani I, AlKhayer A. Visualized malware multi-classification framework using fine-tuned CNN-based transfer learning models. *Appl Sci.* (2021) 11(14):6446. doi: 10.3390/app11146446
42. Almomani I, Alkhayer A, El-Shafai W. An automated vision-based deep learning model for efficient detection of android malware attacks. *IEEE Access.* (2022) 10:2700–20. doi: 10.1109/ACCESS.2022.3140341
43. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. *{TensorFlow}: A System for {Large-Scale} Machine Learning.* Berkeley, CA: USENIX Association (2016). p. 265–83.
44. Brownlee J. *Deep learning with python: develop deep learning models on theano and TensorFlow using keras.* Auckland: Machine Learn Mastery (2016).
45. Géron A. *Hands-on machine learning with scikit-learn, Keras, and TensorFlow.* O'Reilly Media, Inc (2022).
46. Hodnett M, Wiley JF. *R Deep Learning Essentials: A Step-by-step guide to Building Deep Learning Models Using TensorFlow, Keras, and MXNet.* Birmingham: Packt Publishing Ltd (2018).
47. Joseph FJJ, Nonsiri S, Monsakul A. *Keras and TensorFlow: A Hands-on Experience. Advanced Deep Learning for Engineers and Scientists: A Practical Approach.* Cham: Springer (2021). p. 85–111.
48. Vasilev I, Slater D, Spacagna G, Roelants P, Zocca V. *Python Deep Learning: Exploring Deep Learning Techniques and Neural Network Architectures with Pytorch, Keras, and TensorFlow.* Birmingham: Packt Publishing Ltd (2019).
49. Bera S, Shrivastava VK. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. *Int J Remote Sens.* (2020) 41(7):2664–83. doi: 10.1080/01431161.2019.1694725
50. Ramezanzade S, Laurentiu T, Bakhshandah A, Ibragimov B, Kvist T, Björndal L. The efficiency of artificial intelligence methods for finding radiographic features in different endodontic treatments—a systematic review. *Acta Odontol Scand.* (2023) 81(6):422–35. doi: 10.1080/00016357.2022.2158929
51. Li Y, Zeng G, Zhang Y, Wang J, Jin Q, Sun L, et al. AGMB-transformer: anatomy-guided multi-branch transformer network for automated evaluation of root canal therapy. *IEEE J Biomed Health Inform.* (2021) 26(4):1684–95. doi: 10.1109/JBHI.2021.3129245
52. Mulryan-Kyne C. Teaching large classes at college and university level: challenges and opportunities. *Teach Higher Educ.* (2010) 15(2):175–85. doi: 10.1080/13562511003620001
53. Qualtrough A. Undergraduate endodontic education: what are the challenges? *Br Dent J.* (2014) 216(6):361–4. doi: 10.1038/sj.bdj.2014.227
54. Saghir MA, Garcia-Godoy F, Gutmann JL, Lotfi M, Asgar K. The reliability of artificial neural network in locating minor apical foramen: a cadaver study. *J Endod.* (2012) 38(8):1130–4. doi: 10.1016/j.joen.2012.05.004
55. Almutairi M, Alattas MH, Alamoudi A, Bahammam SA, Zidane B, Almutairi N, et al. Challenges assessment in endodontics among undergraduate students. *Cureus.* (2023) 15(8):e43215. doi: 10.7759/cureus.43215
56. Barakat RM, Matoug-Elwerfelli M, Almohareb RA, Balto HA. Influence of preclinical training on root canal treatment technical quality and confidence level of undergraduate dental students. *Int J Dent.* (2021) 2021:1–8. doi: 10.1155/2021/9920280