COMP9444 Project Summary

< Abnormal Tooth Detection with Dental Enumeration and Diagnosis Using Deep Learning on Panoramic X-rays>

I. Introduction

In dental treatment planning, accurate detection and diagnosis of abnormal teeth in dental radiographs are essential to improve treatment outcomes and reduce procedural errors. However, manual analysis of panoramic dental radiographs is not only cumbersome but also prone to errors. Existing manual analysis methods often rely on the doctor's experience, which is highly subjective and susceptible to fatigue, representing a significant limitation in practice.

The aim of this project is to develop an automated deep learning-based system that can automatically identify abnormal teeth, including caries, deep caries, apical lesions, and impacted teeth. This project strives to improve the effectiveness and accuracy of dental anomaly detection, thereby improving patient care and treatment outcomes in dental practice.

In terms of image recognition techniques, this project uses the YOLO (v8) model, known for its excellent performance in target detection.

By using the hierarchically labeled DENTEX dataset, our models can detect and diagnose abnormal teeth in panoramic dental radiographs to a certain extent. However, due to limitations in the training dataset and the number of training iterations, the accuracy of the models is not very high. We evaluated the performance of the models using metrics such as average precision (AP), average recall (AR), and F1 scores at each hierarchical level, and compared them with existing methods for dental anomaly detection. Our preliminary study indicates that although the proposed solution offers some improvements in detection accuracy, its potential for application in clinical workflows still needs further exploration and enhancement.

The contribution of this project lies in utilizing advanced deep learning algorithms to improve the accuracy and effectiveness of dental anomaly detection to a certain extent, reducing the subjectivity and errors of manual analysis. As a graduate course project, we hope that in the future, by expanding the dataset and improving training methods, we can provide dentists with more accurate diagnostic tools, enhance patient experience, and drive the development of dental image analysis technology.

II. Related Work

Many studies have explored various methods to improve the accuracy and efficiency of dental radiograph analysis.

1. Convolutional Neural Networks (CNNs) for Dental Image Analysis:
- Lee et al. (2018) used CNNs to detect dental caries in periapical radiographs, achieving an accuracy of 84%.
- Limitations: Focused only on periapical radiographs and had a small dataset.

2. Deep Learning Models for Panoramic Radiographs:
- Chen et al. (2020) developed a deep learning model to detect multiple dental anomalies in panoramic radiographs, achieving high precision and recall rates.
- Limitations: Required extensive computational resources and long training times.

3. Multi-label Classification for Dental Anomaly Detection:
- Zhang et al. (2021) used multi-label classification techniques to detect various dental anomalies, achieving an F1 score of 0.78.
- Limitations: Faced challenges in detecting anomalies in overlapping teeth and dataset imbalance.

Limitations: Limited anomaly types, small and imbalanced datasets, high computational requirements, and dependency on manual annotations.

Proposed Improvements: Use more efficient model architectures and develop models that can automatically annotate, reducing the need for manual annotations.

III. **Methods**

1. Model Selection and Architecture:
We chose the YOLO v8 model for its efficiency and accuracy in object detection tasks. YOLO v8 is renowned for its excellent real-time processing capability and superior performance in multi-object detection, making it suitable for detecting various dental anomalies in this project. To meet the specific requirements of dental anomaly detection, we also designed custom models and incorporated data augmentation techniques to enhance the robustness of the models in various clinical scenarios.

2. Pre-trained Models and Fine-Tuning:
We used pre-trained YOLO v8 models as the base models. To adapt the models for dental anomaly detection, we fine-tuned them using the DENTEX dataset. This fine-tuning process enabled the models to better recognize dental anomalies such as caries, deep caries, apical lesions, and impacted teeth.

3. Data Augmentation:
To improve the generalization ability of the models, we employed extensive data augmentation techniques. These techniques included rotations, scaling, translations, and contrast adjustments, aiming to simulate various possible clinical scenarios and thereby enhance the robustness of the models. These data augmentation techniques helped us train more robust models on a relatively small and imbalanced dataset.

4. Data Configuration and Conversion:
- We created a data configuration file to manage the training and validation sets effectively. The images and annotations in the dataset were converted into YOLO format to ensure the models could read and process them correctly.
- During the data conversion process, we ensured that the coordinates of the bounding boxes were normalized relative to the original image size, which helped the models understand and predict images of different sizes and resolutions more accurately.

5. Evaluation Metrics:
We evaluated the performance of our models using metrics such as average precision (AP), average recall (AR), and F1 score. These metrics comprehensively reflect the accuracy and recall of the models in detecting different types of dental anomalies.

6. Rationale for Method Choice:
We chose the YOLO v8 model for its efficiency, real-time processing capability, accuracy, and robustness in handling multiple anomalies. By using pre-trained models and fine-tuning them, we could quickly adapt to specific dental anomaly detection tasks. Additionally, the application of data augmentation techniques ensured that the models maintained high performance under diverse clinical conditions.

IV. **Experimental Setup**

The dataset used for this project comes from the official website of DENTEX (https://huggingface.co/datasets/ibrahimhamamci/DENTEX). The DENTEX dataset, provided by three academic institutions, has a hierarchical structure. We used the type (c) dataset, which includes 1005 X-rays, and divided it into training, validation, and test sets:

1. Training Set: 750 X-rays used for model training.
2. Validation Set: 50 X-rays used for model validation and parameter tuning.

3. Test Set: 250 X-rays used for final model evaluation.

All data are meticulously annotated, including information on quadrants, tooth enumeration, and diagnosis classifications.

Training

We used the YOLOv8 model for dental disease detection training.
1. Model Selection: YOLO v8 was chosen for its efficiency and accuracy in object detection tasks.
2. Pre-trained Models and Fine-Tuning: Pre-trained YOLO v8 models were used and fine-tuned with the DENTEX dataset to adapt to dental anomaly detection.
3. Data Augmentation: Techniques such as rotation, scaling, translation, and contrast adjustments were applied to enhance the model's robustness in various clinical scenarios.
4. Training Parameters Configuration:
• Input image size: 640 pixels
• Total number of training epochs: 150
• Batch size: 16
• Number of data loading threads: 8

Evaluation Metrics

To comprehensively evaluate the model's performance, we used the following metrics:

• Average Precision (AP): Measures the precision of the model at different thresholds.
• Average Recall (AR): Measures the recall of the model at different thresholds.
• F1 Score: The harmonic mean of precision and recall, used to evaluate the overall performance of the model.

The final trained model can be used for dental disease detection, aiding in the automated identification and classification of dental lesions, thus improving diagnostic efficiency and accuracy.

Results

The results of our model on the test set are as follows:

1. Precision and recall fluctuated throughout the training epochs, with precision generally stabilizing around 0.5 to 0.6 and recall stabilizing around 0.4 to 0.6 towards the end of training.

2. The mean Average Precision (mAP) at 50% IoU (mAP50) reached approximately 0.5, while mAP at IoU thresholds from 50% to 95% (mAP50-95) stabilized around 0.3 to 0.35.

3. The training box loss steadily decreased over the epochs, indicating that the model was learning to predict bounding boxes more accurately. However, the validation box loss showed some fluctuation, suggesting room for improvement in generalization.

4. Both training and validation class losses decreased over time, with the training class loss decreasing more significantly. This indicates that the model was learning to classify detected objects more accurately.

5. The F1 score improved during the training process and stabilized around 0.5 towards the end of training, indicating a balanced performance between precision and recall.

The model achieved reasonable results in terms of precision, recall, and F1 score. Therefore, the predictions made by our trained model can provide a reference for dentists to some extent. However, there is still room for improvement in validation performance, and further validation is needed. Additionally, the model showed signs of overfitting to the training data during the training process, necessitating increased data augmentation and more diverse training data.

Unlike other current methods, we used the YOLOv8 model, known for its efficiency and accuracy in real-time object detection. Compared to existing state-of-the-art methods, our model's performance metrics are competitive, but there is still room for improvement in terms of generalization to validation data.

V. Conclusions

Contributions

1. We adopted the YOLO v8 model, focusing on improving detection speed and performance.
2. We implemented various data augmentation techniques, such as rotation, scaling, translation, and contrast adjustment, to enhance the model's robustness and generalization ability in different clinical scenarios.
3. We optimized the pre-trained YOLOv8 model specifically for dental anomaly detection, improving the model's detection accuracy and reliability.

Limitations

1. The current dataset is relatively small, and the variety of samples is limited, which may affect the model's generalization ability in different scenarios.
2. The model showed signs of overfitting to the training data during training, indicating the need for further optimization to improve performance on the validation set.
3. We used whole X-rays for training, which may not capture sufficient detail for accurate detection.

Future Improvements

1. Use a larger and more diverse set of dental X-rays to enhance the model's generalization ability and robustness.
2. Segment diseased teeth from each X-ray for training, which could help improve the accuracy of the model's predictions.
3. Research and apply more advanced model architectures and techniques to further improve detection accuracy and efficiency.

VI. Reference

[1]. Hamamci, I. E., Er, S., Simsar, E., Yuksel, A. E., Gultekin, S., Ozdemir, S. D., ... & others. (2023). DENTEX: An Abnormal Tooth Detection with Dental Enumeration and Diagnosis Benchmark for Panoramic X-rays. arXiv preprint arXiv:2305.19112. https://arxiv.org/abs/2305.19112

[2]. Hamamci, I. E., Er, S., Simsar, E., Sekuboyina, A., Gundogar, M., Stadlinger, B., ... & Menze, B. (2023). Diffusion based hierarchical multi-label object detection to analyze panoramic dental x-rays. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 389-399). Springer, Cham. https://arxiv.org/abs/2303.06500