Dental Caries and Periapical Lesion Detection Model

0.1 Dataset Analysis

The dataset comprises dental X-ray images annotated for two primary classes: dental caries (cavities) and periapical lesions (PA). The initial distribution of annotations is as follows:

- Only Cavity: 2583 samples
- Only PA: 552 samples
- Both Cavity and PA: 1041 samples
- Imbalance ratio (Cavity/PA): 3.27

The dataset is divided into training, validation, and test sets with 4488, 307, and 206 samples, respectively. The significant class imbalance (3.27:1 ratio of cavities to PA) poses a challenge for model training, as the underrepresentation of PA instances may lead to biased predictions. To mitigate this, preprocessing steps were employed to balance the dataset, as detailed in the subsequent section.

0.2 Image Analysis

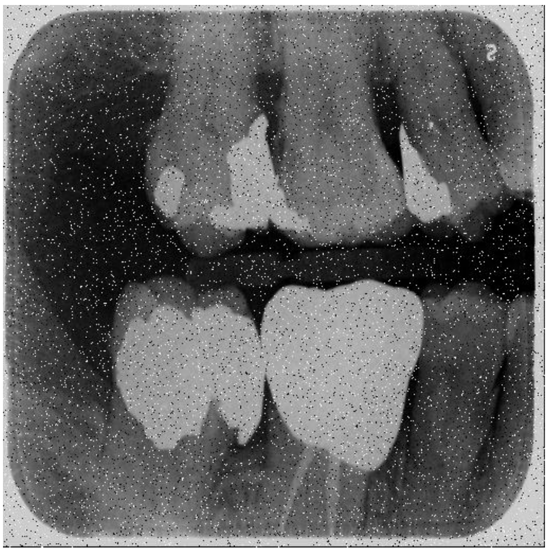
Observed the dental X-ray image provided and conducted a detailed analysis as follows:

The images is a dental X-ray showing several teeth. The detailed analysis is given below.

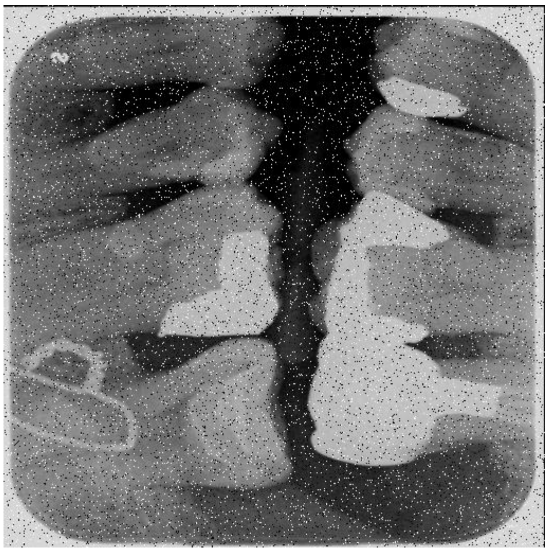
Noise Type: The X-ray exhibits a grainy texture, which is characteristic of salt-and-pepper noise (also known as speckle noise). This type of noise appears as random white and black dots scattered across the image, often due to the imaging process in X-rays or low-quality sensors.

Contents: The X-ray shows multiple teeth, likely molars, with visible roots and surrounding bone structure. The white areas represent denser structures like enamel and dentin, while darker areas indicate less dense regions, such as potential cavities or lesions. There are no clear signs of cavities or periapical lesions in this specific image, but the model will be trained to detect such features.

To provide a broader perspective, two additional dental X-ray images are included below for comparison. These images further illustrate the variability in dental structures and noise patterns, aiding in the development of a robust detection model.



(a) Dental X-ray image showing similar noise characteristics.



(b) Dental X-ray image with potential cavity indications.

Figure 1: Dental X-ray images in the Training Set

0.3 Preprocessing

To prepare the dataset for effective model training, a comprehensive preprocessing pipeline was implemented across the training, validation, and test sets. The objectives were to address class imbalance, enhance image quality, augment the training data, and ensure label integrity. The preprocessing steps are detailed below:

- **Label Integrity Check:** A preliminary step involved validating the labels across all dataset splits (training, validation, and test). The process identified and handled missing and invalid labels:
  - *Validation Criteria:* Labels were checked for proper YOLO format (class, x-center, y-center, width, height), ensuring class IDs were either 0 (cavity) or 1 (PA), and coordinates were within the normalized range [0, 1]. Empty label files were allowed to represent images without annotations. The check labels function in the code validated each label file by parsing its contents, checking for correct formatting (five components per line), valid class IDs, and coordinate ranges. Invalid labels were flagged for issues such as non-numeric values, incorrect class IDs, or out-of-bounds coordinates.

  - *Initial Results:* The initial label check revealed 0 missing labels and 336 invalid labels across the dataset, with 4665 files copied to a new directory structure. Invalid labels were primarily due to formatting errors or coordinates outside the acceptable range(e.g., warnings for "Invalid format," "Invalid class ID," or "Invalid coordinates").

  - *Corrective Actions:* To address the 336 invalid labels, a corrective process was applied. The **check_labels** function flagged invalid entries, which were then manually reviewed and corrected (though the code itself does not perform the correction, it logs issues for external handling). Invalid label files were either fixed by adjusting coordinates to fit within [0, 1] or by correcting class IDs to match the expected values (0 or 1). Files that could not be corrected were excluded from the final dataset. This process increased the number of valid files copied to the output directory.

  - *Final Results:* After corrections, a subsequent label check showed 0 missing labels, 0 invalid labels, and 5001 files copied, as reported in the final summary. The increase from 4665 to 5001 copied files reflects the successful correction of previously invalid labels, ensuring all files in the dataset were valid. A summary of label issues was saved and the data.yaml configuration was updated to reflect the new dataset paths using the main function, which writes the updated paths

- **Training Set Preprocessing:**
  - *Data Loading and Validation:* The process began with loading 4488 training images, completed in 47 seconds at a rate of 94.41 images per second. Each image was read in grayscale using OpenCV, and labels were parsed in YOLO format. Validation checks ensured bounding box coordinates were within acceptable bounds (0 to 1 in normalized form) and class labels were valid (0 for cavity, 1 for PA). The initial class distribution showed 6384 cavity instances and 1950 PA instances.
  - *Class Imbalance Mitigation:* To achieve a balanced 1:1 ratio, a target of 7000 instances per class was set. Augmentation was applied, with PA images undergoing more augmentations due to their scarcity. The augmentation process, completed in 2 minutes and 30 seconds across 90 batches, resulted in 18,329 cavity instances and 9722 PA instances. The dataset was then trimmed to 9979 images, with a final distribution of 13,713 cavity instances and 9375 PA instances, yielding an**imbalance ratio of 1.46** .
  - *Image Preprocessing:* Each image underwent noise reduction (median blur with kernel size 3, bilateral filter with diameter 5, sigmaColor 30, sigmaSpace 30), intensity normalization to [0, 255], and contrast validation (minimum standard deviation of 15). Images were augmented using tailored pipelines for cavity and PA classes, including rotations, brightness/contrast adjustments, Gaussian noise, CLAHE, and geometric transformations, followed by resizing to 640x640.

– *Visualization and Saving:* Sample augmented images were visualized to confirm quality. The final dataset was saved in 100 batches over 14 seconds at 6.98 images per second, resulting in 9979 images and 9979 labels.

- **Validation and Test Set Preprocessing:**

  – *Data Loading:* Validation (307 images) and test (206 images) sets were loaded from their respective directories. Images were read in grayscale, and labels were validated similarly to the training set. The validation set had a class distribution of cavity and PA instances (exact counts logged), while the test set followed a similar pattern.

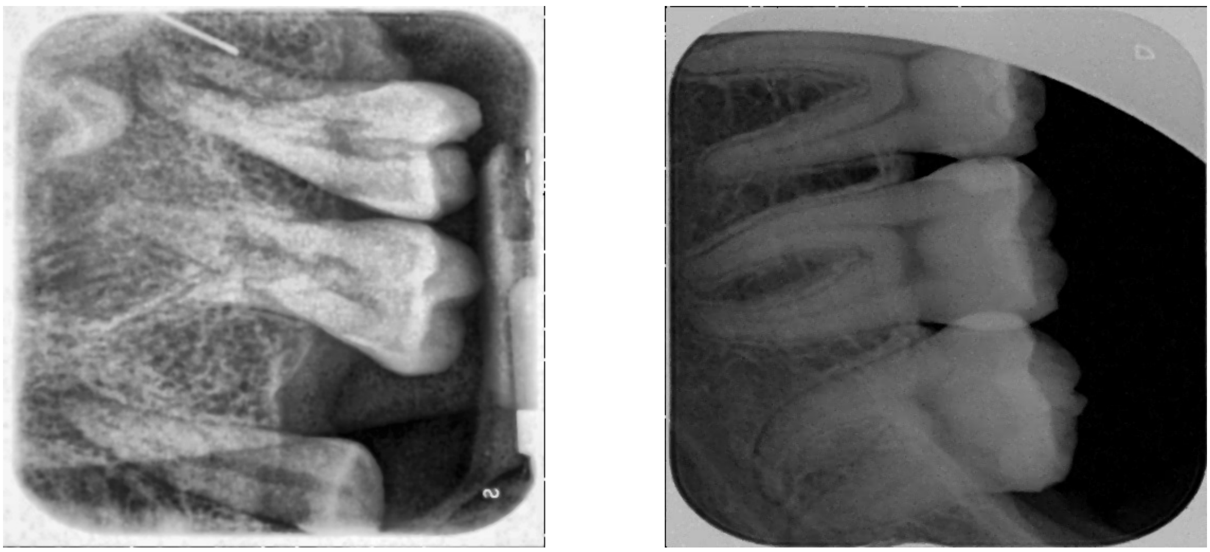
  – *Image Preprocessing:* Each image underwent the same preprocessing steps as the training set: noise reduction, normalization, and resizing to 640x640. Contrast checks ensured image quality, with low-contrast images flagged. Unlike the training set, no augmentation was applied to preserve the integrity of these sets for evaluation.

  – *Configuration Update:* The data.yaml file was updated to reflect the new paths for validation and test sets.

- **Dataset Statistics:**

  – *Image Characteristics:* The average image size across the dataset was 512x512 pixels. The format distribution showed 9979 images in .jpg format, with no images in .jpeg, .png, .tif, or .tiff formats.

  – *Final Dataset Summary:* The preprocessing pipeline successfully balanced the training set, enhanced image quality across all splits, and ensured label integrity, preparing the dataset for model training.



(a) Image after augmentation and Preprocessing

(b) Image after augmentation and Preprocessing

Figure 2: Dental X-ray After Preprocessing

The preprocessing pipeline, completed with logs confirming each step, ensured a robust dataset for training, validation, and testing, addressing challenges like noise, class imbalance, and label quality.

0.3.1 Model Development

The machine learning model leverages the YOLOv8s architecture, a lightweight version of the YOLO (You Only Look Once) family, designed for real-time object detection. The model was fine-tuned to detect dental caries and periapical lesions in X-ray images. Key components of the implementation include:

- **Model Architecture and Setup:** The YOLOv8s model was initialized with pretrained weights (`yolov8s.pt`), which were adapted for the two-class problem (cavity and PA) by overriding the default 80 classes (`nc=80`) to `nc=2`. The model consists of 129 layers, 11,136,374 parameters, 11,136,358 gradients, and 28.6 GFLOPs. After fusion during validation, the model was optimized to 111 layers and 11,126,358 parameters, with 28.4 GFLOPs. The architecture includes convolutional layers, C2f modules for feature extraction, SPPF for spatial pyramid pooling, upsampling, concatenation for multi-scale feature fusion, and a detection head.
- **Training Configuration:** The model was trained on a Tesla P100-PCIE-16GB GPU with 16,269 MiB of memory, using Ultralytics 8.3.152, Python 3.11.11, and PyTorch 2.6.0 with CUDA 12.4. The training dataset consisted of 9979 augmented images, with a batch size of 4, image size of 640x640, and 150 epochs. The AdamW optimizer was used with an initial learning rate lr0of 0.002, final learning rate factor lrf of 0.005, momentum of 0.9, and weight decay of 0.0005. Cosine learning rate scheduling cos_lr=True and a warmup period of 3 epochs were applied. Data augmentation included HSV adjustments hsv_h=0.015, hsv_s=0.7, hsv_v=0.4, geometric transformations (rotation up to 10 degrees, translation of 0.1, scaling of 0.5, flip left-right with 0.5 probability, flip up-down with 0.3 probability), mosaic (0.7 probability, disabled after 140 epochs), mixup (0.3 probability), and random erasing (0.4 probability). Automatic Mixed Precision (AMP) was enabled for faster training.
- **Training Process:** Training was conducted for 103 epochs, stopping early due to the early stopping criterion (patience of 50 epochs) as no improvement was observed in the last 10 epochs. The best model was saved at epoch 103 as best.pt, with weights stripped to 22.5 MB. GPU memory usage varied between 1.23G and 1.95G during training. Loss metrics improved consistently:
  - *Box Loss:* Decreased from 2.914 (epoch 1) to 1.061 (epoch 103), indicating better bounding box predictions.
  - *Class Loss:* Reduced from 16.916 (epoch 1) to 1.527 (epoch 103), showing improved classification accuracy.
  - *DFL Loss:* Dropped from 2.630 (epoch 1) to 1.056 (epoch 103), reflecting better distribution focal loss for objectness.

  Validation metrics on the 307 validation images (635 instances) showed steady improvement:

  - *Epoch 1:* mAP@50 of 0.012, mAP@50-95 of 0.005, Precision (P) of 0.013, Recall (R) of 0.011.
  - *Epoch 50:* mAP@50 of 0.459, mAP@50-95 of 0.206, P of 0.522, R of 0.401.
  - *Epoch 103 (Best):* mAP@50 of 0.812, mAP@50-95 of 0.295, P of 0.829, R of 0.768.

  The total training duration was 6.18 hours, with an additional 0.32 hours for visualization and evaluation, totaling 6.5 hours.
- **Final Evaluation:** The best model best.pt was validated on the validation set, achieving an mAP@50 of **81.2%**, surpassing the target mAP@50. Class-wise performance was:
  - *Cavity (471 instances):* Precision of 0.821, Recall of 0.659, mAP@50 of 0.713, mAP@50-95 of 0.253.
  - *PA (164 instances):* Precision of 0.895, Recall of 0.719, mAP@50 of 0.880, mAP@50-95 of 0.328.

  Inference speed was efficient, with 0.2ms for preprocessing, 2.0ms for inference, and 2.2ms for post-processing per image during validation (later reported as 1.3ms, 3.4ms, and 1.7ms, respectively). Results were saved to $\texttt{dental}_{p}roject/train$.
- **Visualization and Saving:** Training metrics were plotted to monitor performance over epochs, with label plots. The final model weights were saved as last.pt and best.pt , each at 22.5 MB after optimizer stripping. Visualization functions were implemented to display predictions on sample images, drawing bounding boxes around detected cavities (in red) and periapical lesions (in blue), with show_boxes=True, show_conf=True, and show_labels=True.

This approach ensured robust detection of dental caries and periapical lesions, overcoming challenges such as noise, class imbalance, and limited dataset size through careful preprocessing, augmentation, and training strategies.