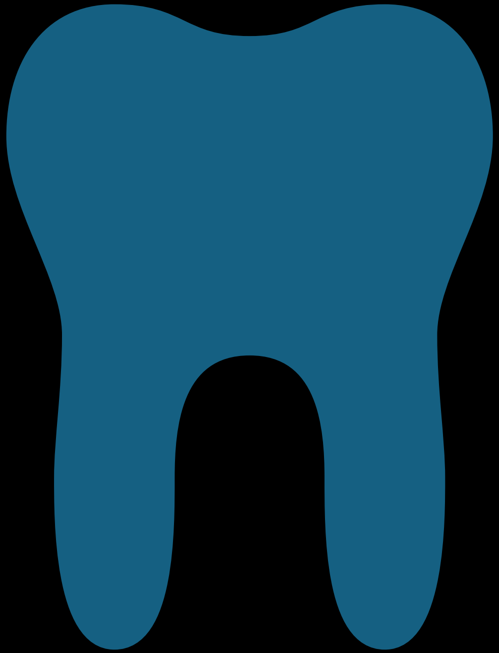
Hierarchical Multi-Label Object Detection to Analyze Panoramic Dental X-rays



Amruta Parulekar (20D070009)

Annie D'Souza(20D070028)

Bhavya Kohli(20D070021)

Bhavya Singh(200040036)

Sanjhi Priya(20D070070)

Tejaswee Sulekh(20D070082)

PROBLEM STATEMENT

Interpreting panoramic X-rays in dental radiology is time-consuming and prone to misdiagnosis due to work exhaustion.

Challenges Faced

- **Variations in Anatomy**
- **Insufficient publicly available annotated data**

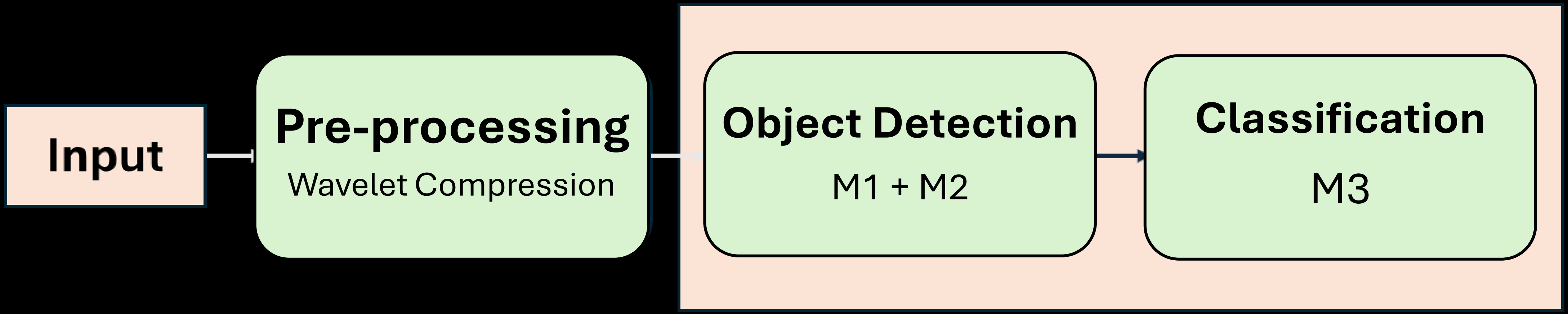
Objective

Develop algorithms for accurate detection of abnormal teeth and associated diagnosis to enhance treatment

Goal

- **Improve computation time and efficiency using techniques like wavelet transforms.**
- **Address class imbalance using advanced loss functions for classification.**
- **Experiment with object detection techniques to surpass baseline performance.**

PIPELINE

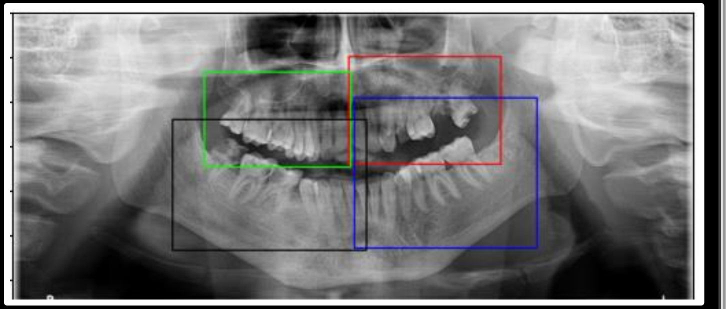


The main model was implemented by first performing **object detection** followed by **classification**:

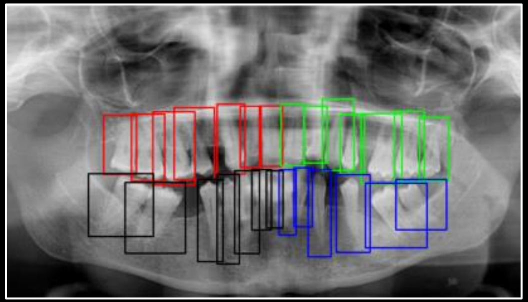
- ❑ Detect all teeth first
- ❑ Extract teeth patches and classify individual teeth as *good* or *diseased*.

DATASET

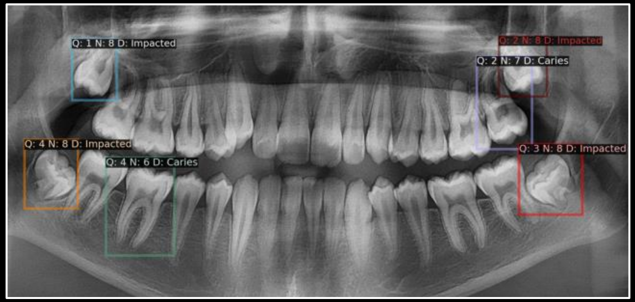
- 3 datasets, each with different images and different labels



- **Dataset 1:**
  - Contains bounding boxes for only quadrants
  - Bounding boxes have quadrant labels

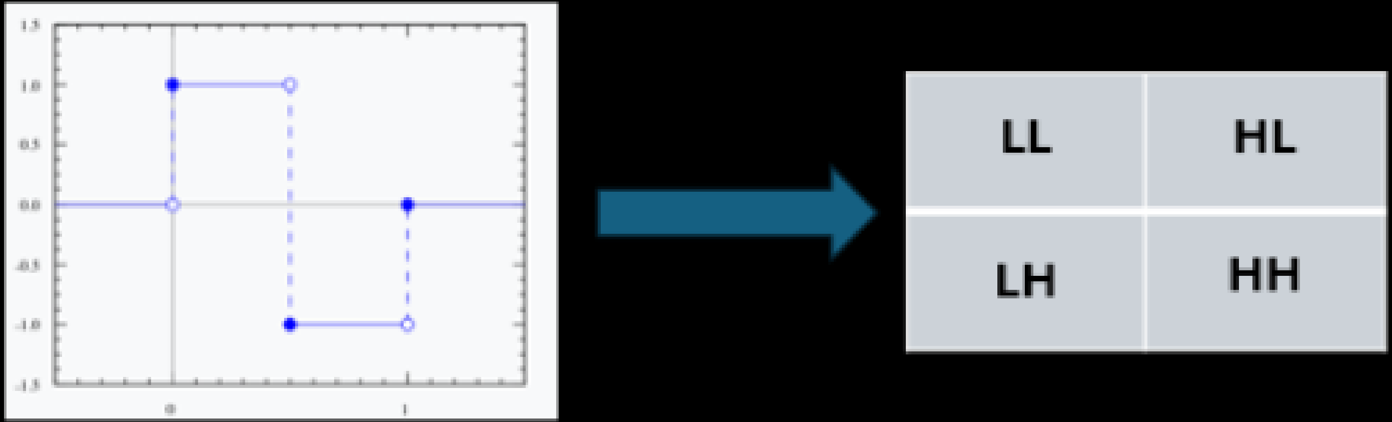


- **Dataset 2:**
  - Contains bounding boxes for only enumerations
  - Bounding boxes have quadrant and enumeration labels



- **Dataset 3:**
  - Contains bounding boxes for only diseased teeth
  - Bounding boxes for normal teeth are not available
  - Bounding boxes have quadrant, enumeration, and disease labels

PRE-PROCESSING: WAVELET COMPRESSION



HAAR Decomposition of the image produces these components, where

H = Highpass

L = Lowpass

Finding the correct weights:

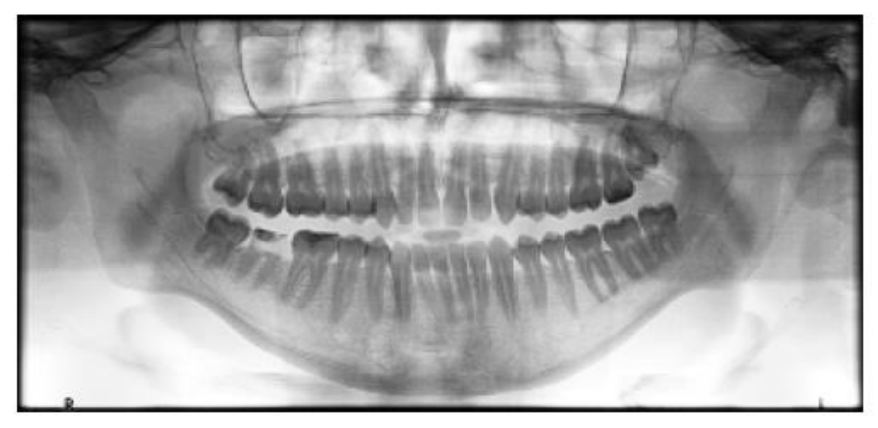
1. **Grid search** over all possibilities, keeping the weight ratio (M : N : N : N)
2. Calculating the **MSE** between the original and compressed image
3. **Interpolated** using different transformation such as
   1. INTER_LINEAR
   2. INTER_AREA
   3. INTER_CUBIC

M	Error (MSE)
0.1	483115.2
0.228	5663.8
0.357	**2323.2**
0.485	2626.9
0.613	3055.6
0.742	3453.3
0.87	3773.7

PRE-PROCESSING: WAVELET COMPRESSION

- Sample compressed images:

Original : (1316, 2765)



3-Times : (165, 346)

5-Times : (42, 87)

OBJECT DETECTION MODELS M1 & M2

- Two models—M1 and M2—used for quadrant and tooth detection respectively
- DETRs with Collaborative Hybrid Assignments (Co-DETR), the current SOTA on object detection benchmarks, is the choice of architecture for both models
- Rationale behind this dual-model setup:
  - We do not need all images to have all types of annotations
  - Allows use of smaller, more specialized models instead of a single large model with low interpretability
  - Can use targeted human supervision to finetune specific portions of the workflow instead of re-training the whole (very deep) model from scratch every time

PROCESSING BETWEEN MODELS

- **WHY?**
  - Third dataset **only** contains bounding boxes for **diseased teeth** so without this pipeline, we can't **obtain bounding boxes for normal teeth**
- **For training M2:**
  - Second dataset →M1→ Quadrant bounding boxes
  - **Crop** images to get 4 quadrants and **reflect** quadrant 2,3,4 to align with quadrant 1
  - For the **enumeration bounding box labels** in annotations file:
    - **Shift** the bounding boxes according to quadrant positions
    - **Reflect** the bounding box **coordinates** and update them to the **required format**.
- **For training M3:**
  - Third dataset →M1→ Quadrant bounding boxes → Quadrant images w/ shifted bbox labels
  - Quadrant images →M2→ Enumeration bounding boxes
    - For bounding boxes with **valid confidence scores**, crop to obtain patches
    - Go through the annotations file and for existing **teeth w/o labels**, assign **normal**

RESULTS OF COMPRESSION

Model	Compression Degree	AP IoU 0.50:0.95	AP IoU 0.75	AR IoU 0.50:0.95
M1	0	0.713	0.907	0.797
	3	**0.713**	**0.890**	**0.795**
	4	0.699	0.867	0.791
	5	0.698	0.900	0.776
Quad	SOTA (2020)	0.651	0.524	0.727
M2	0	0.543	0.583	0.718
	3	**0.542**	**0.587**	**0.727**
	4	0.343	0.275	0.650
Enum	SOTA (2020)	0.494	0.394	0.668

- For M1, compression gives comparative results with slight decrease in each degree.
- For M2, we find compression degree 3 to be optimal, degree 4 suffers greatly.

CLASSIFIER WITH LONG-TAIL SOLUTIONS

TOOTH PATCHES

MODEL: EFFICIENTNETB0
OPTIMIZER: ADAM
LR SCHEDULER: STEPLR
EARLY STOPPING: 6 EPOCHS
CRITERION: CATEGORICAL CROSS ENTROPY

Carries - 604

Deep Carries - 2189

Periapical Lesions - 158

Impacted Teeth – 578

Normal - 16595

CLASS IMBALANCE

Focal Loss

- Alpha=[3,1,13,3], Gamma=2
- Addresses class imbalance by emphasizing difficult-to-classify examples from less represented classes during training.

Intelligent Data subset selection

- Clustered normal class patches – 2000 clusters
- Pick patch closest to centroid as representative
- 16000 images reduced to 2000 representative images

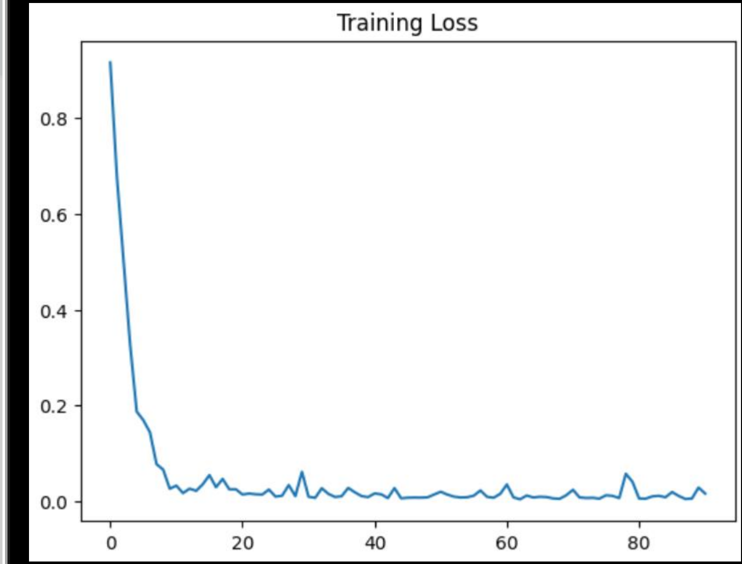
Geometric Augmentations

- Random flips and rotations of rarer classes during training
- Done such that every class has approximately 2000 net images

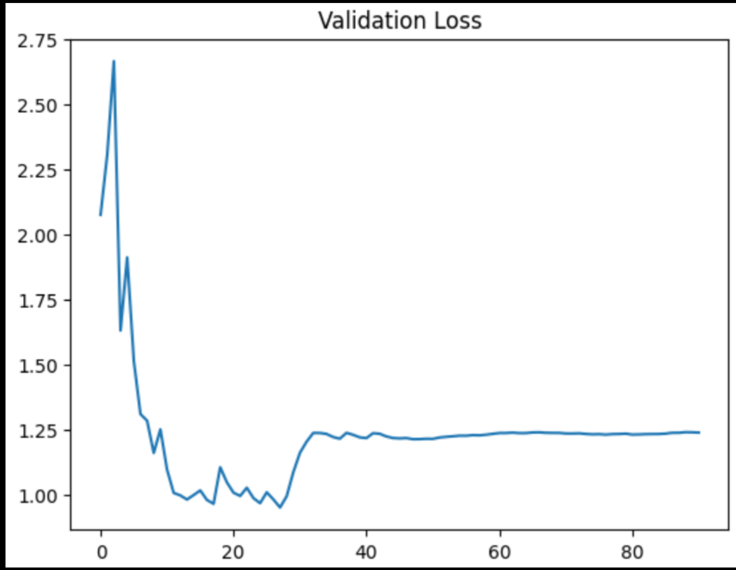
FEATURE EXTRACTION WITH WAVELETS

- Extracted features using the pre final layer of the efficient-net model (1028)
- Concatenated them with:
  - flattened wavelet compressed features
  - flattened wavelet compressed features passes through a linear layer
- Passed this concatenated vector through the final efficient-net layer and trained
- Compression ratios used = 3, 4
- Linear layer divisions used:
  - For R=3: 9 (3196->348)
  - For R=4: 3 (784->261)

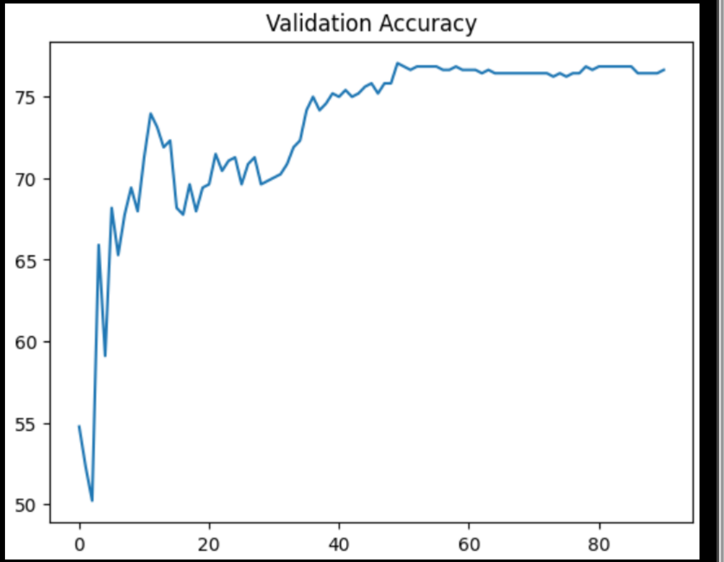
LOSS PLOTS



Training Loss
0.0
0.2
0.4
0.6
0.8
0
20
40
60
80



Validation Loss
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
0
20
40
60
80



Validation Accuracy
50
55
60
65
70
75
0
20
40
60
80

RESULTS

Expt	Class-wise Accuracy(%)					Overall
Baseline	83.22	75.44	8.54	31.7	90.66	72.7%
Geometric augmentations	80.95	83.50	28.57	63.74	82.26	75.6%
Focal Loss	85.39	70.85	41.67	42.65	86.04	74.88%
Wavelets compression degree 3 w GA	91.35	80.24	24.78	34.59	87.32	78.9%
Wavelets compression degree 4 w GA	89.87	78.18	31.82	30.34	86.22	76.0%
(Number of Images)	604	2189	158	578	2000	5529

CONTRIBUTIONS

- **Literature survey** - Everyone
- **Wavelet compression for preprocessing** - Tejaswee
- **M1, M2 Setup**– Bhavya K., Annie
- **Data Processing b/w models** – Bhavya S.
- **M1, M2 experiments** – The Bhavyas
- **M3 classifier + class imbalance correction** – Amruta, Sanjhi
- **Wavelets for classifier** - Annie
- **M3 experiments** – Sanjhi, Amruta, Annie

THANK YOU!

Questions?