



Multi-Objective Processing of Dental Panoramic Radiographs

Ting Lin,¹ Langston Nashold,¹ Poojan Pandya¹

¹Computer Science Department, Stanford University

Stanford
Computer Science

Background

- Dentists diagnose based on panoramic X-Rays
- Time consuming and laborious process to count teeth and look for abnormalities
- Can we use AI to automate it?

Problem

- **Problem:** Given a panoramic radiograph, produce a semantic segmentation teeth mask and a binary label denoting if it has an abnormality
- **Approach:** DeepLabv3 Spatial pooling pyramid, with fusion and multi-task labelling
- **Metrics:** IOU, F1, Accuracy

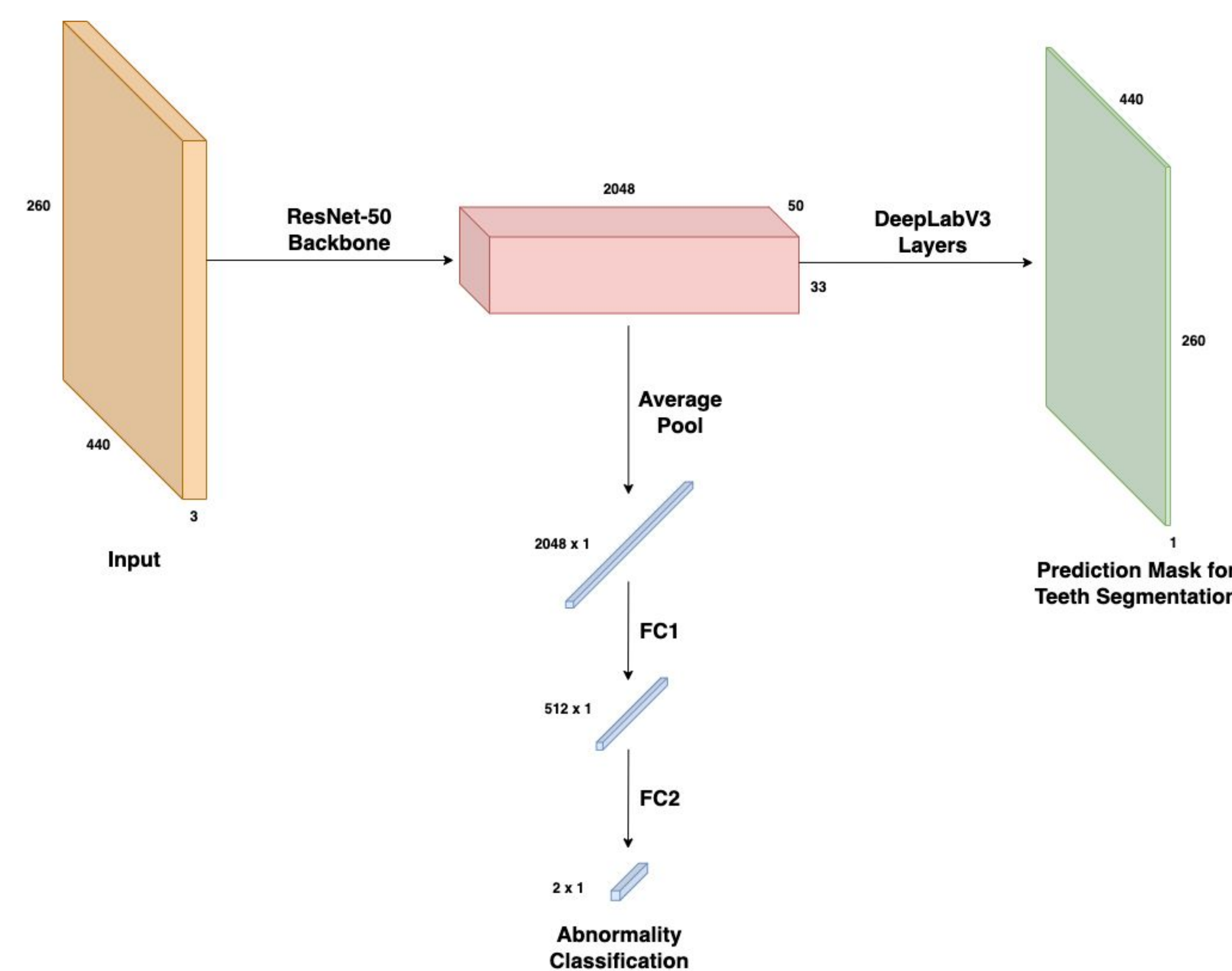
Data

- Tufts Dental Dataset (April 2022)
- Largest publicly available dental dataset
- 1000 panoramic dental radiographs with teeth mask, abnormality mask, and gaze map

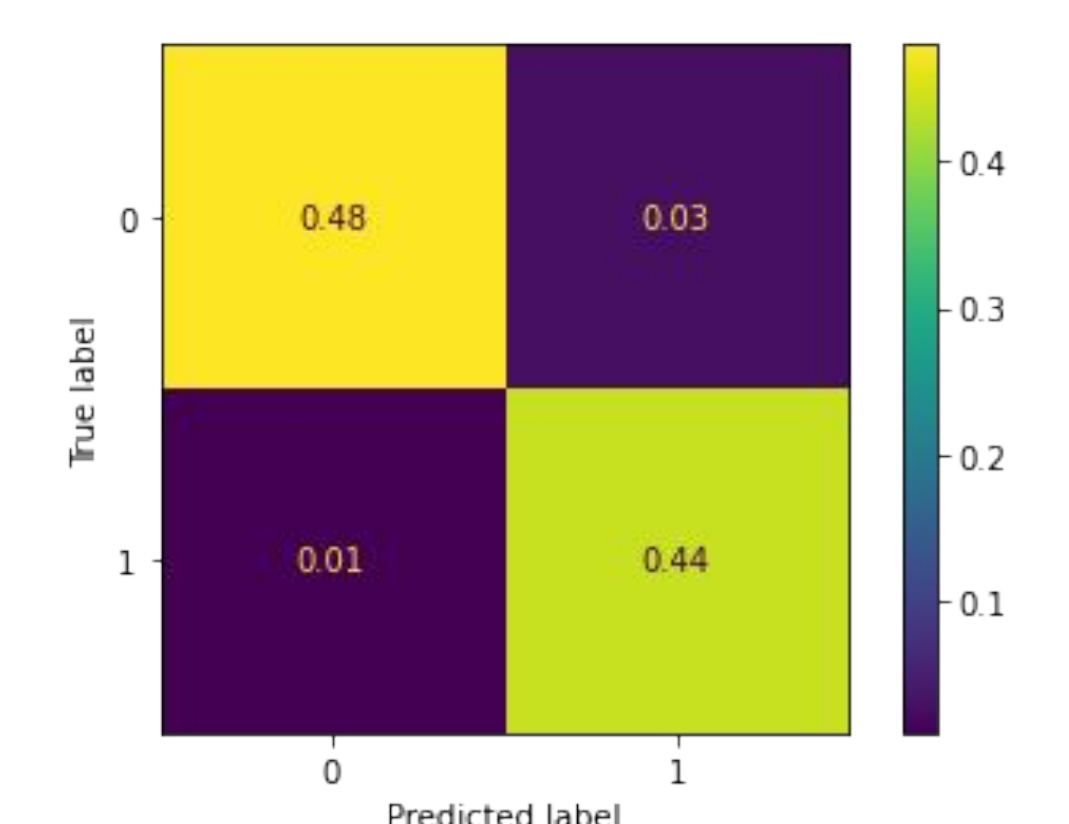


Methods

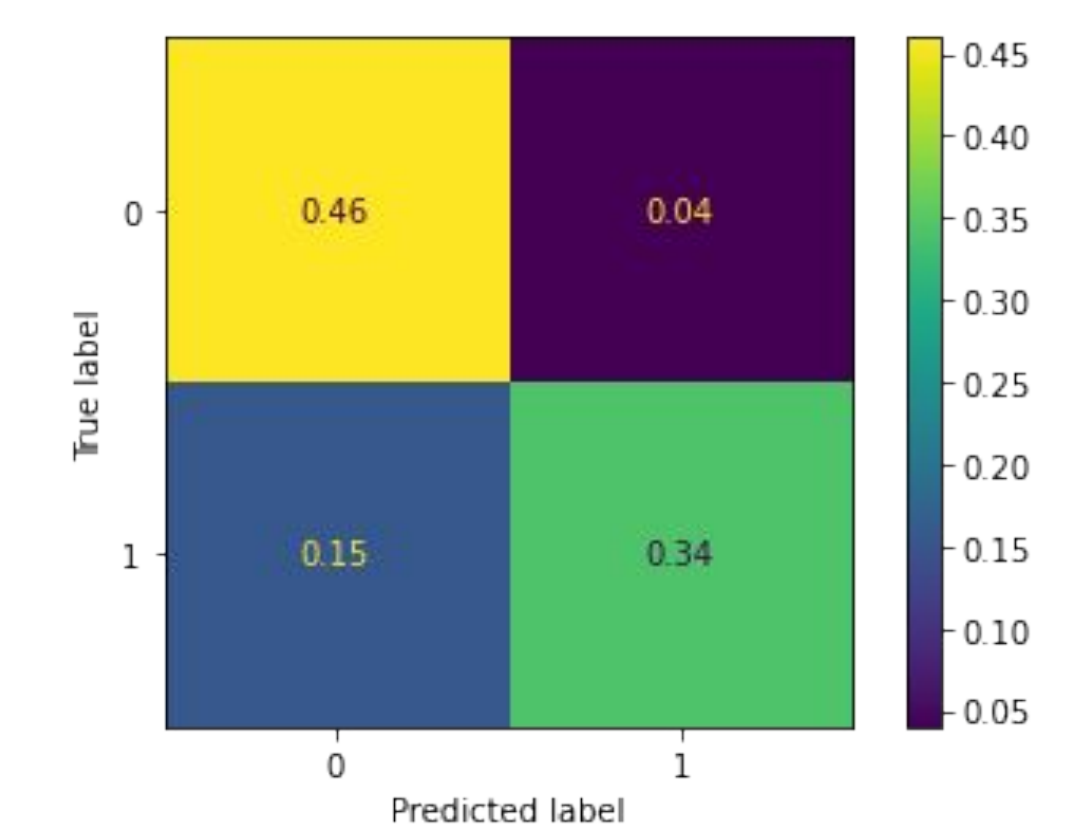
Model 1: Multi-Headed CNN Model



- ResNet50 backbone
- Atrous Spatial Pooling Pyramid on top of backbone to predict teeth mask
- FCN on backbone to do abnormality classification



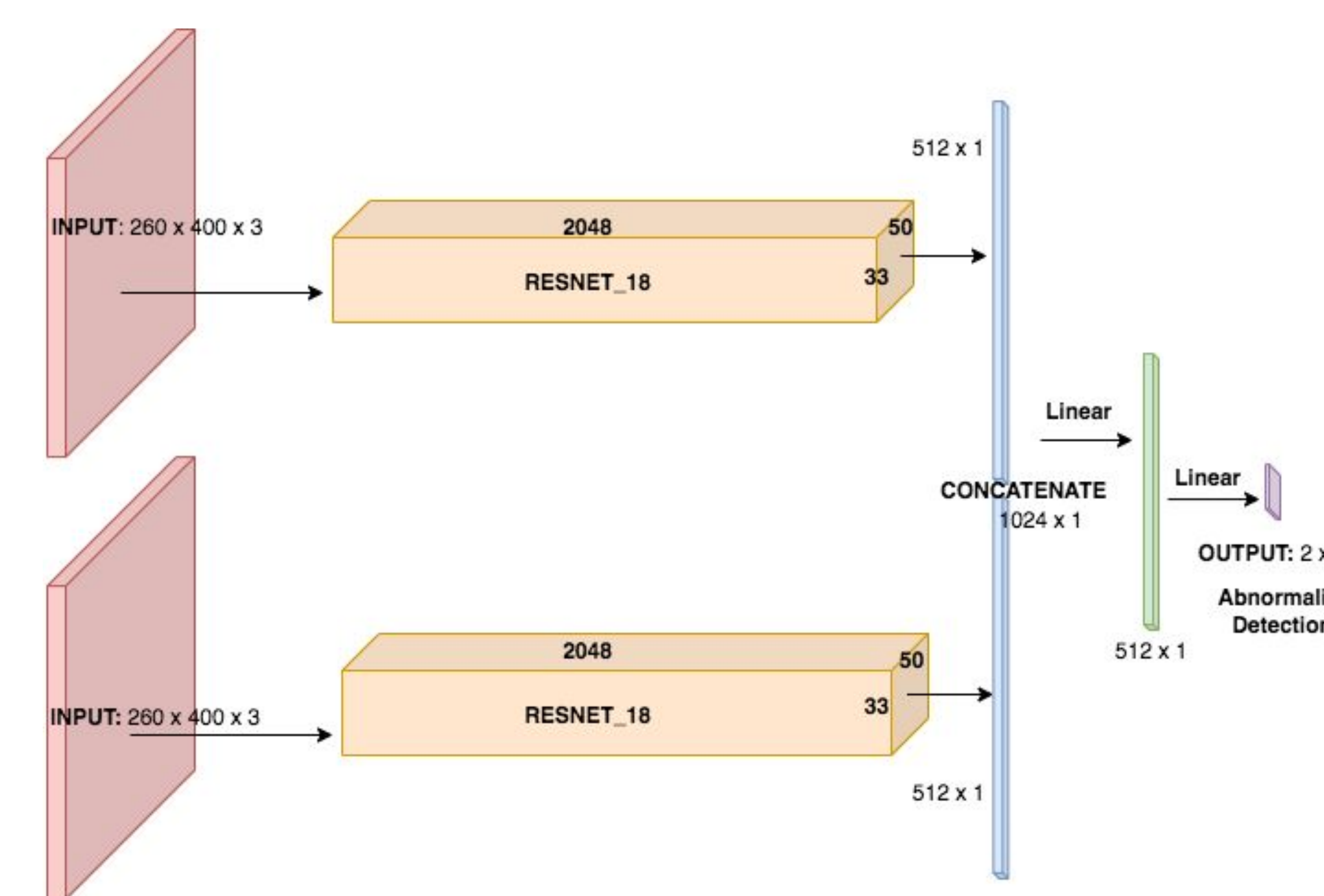
MH Confusion Matrix



FS Confusion Matrix

- Ablation performed on Jaccard, Dice, and BCE Loss

Model 2: Fusion CNN Model



- Gaze map and radiograph are passed into two ResNet-18s
- ResNet output is fused and passed through a two layer FCN to produce abnormality classification

$$\text{Jaccard}(U, V) = \frac{|U \cap V|}{|U \cup V|}$$

$$\text{Dice}(U, V) = \frac{2|U \cap V|}{|U| + |V|}$$

$$\text{BCE}(y, \hat{y}) = -w * (y \log(p) + (1 - y) \log(1 - p))$$

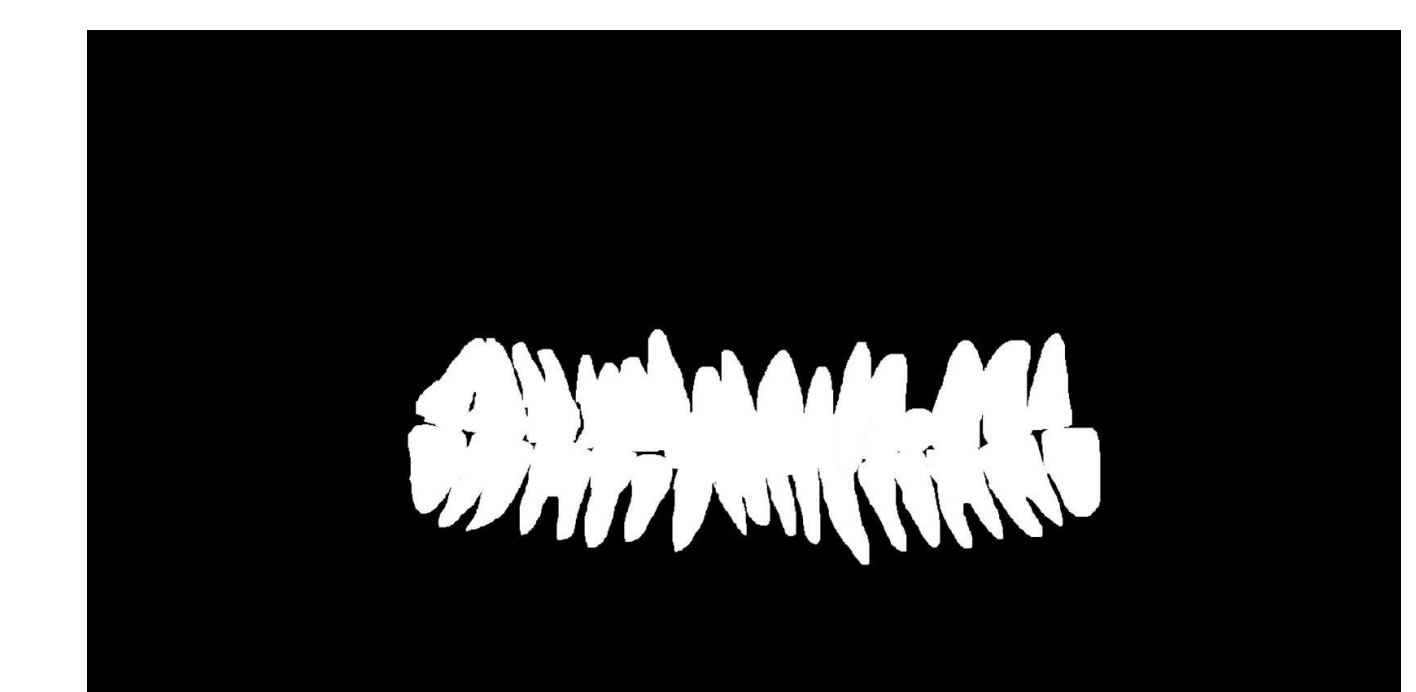
	W-BCE	IOU	Dice
IOU	0.914	0.918	0.918
Accuracy	0.952	0.952	0.951
F1	0.696	0.767	0.757

Results

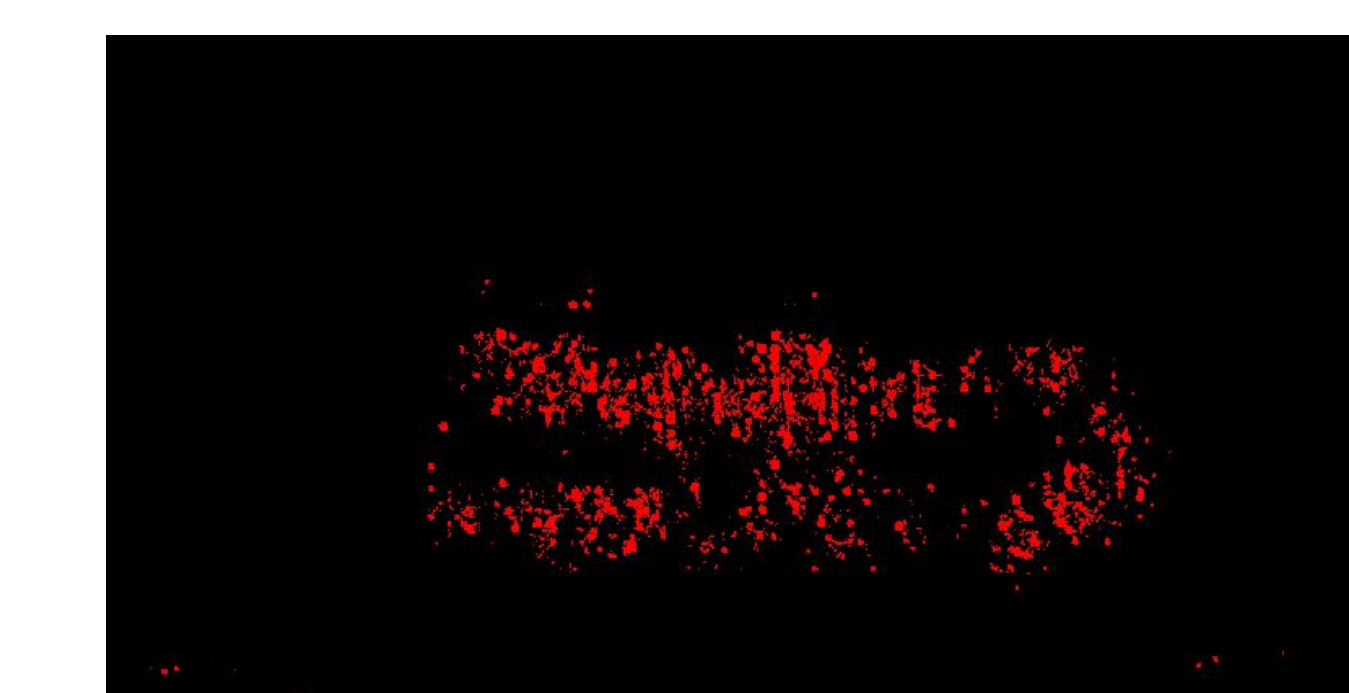
	Segmentation			Detection	
	F1	IOU	Accuracy	F1	Accuracy
Baseline	70.5	91.2	94.9	0.84	0.84
Multi-Headed	76.0	92.4	95.6	0.92	0.92
Fusion	N/A	N/A	N/A	0.69	0.80



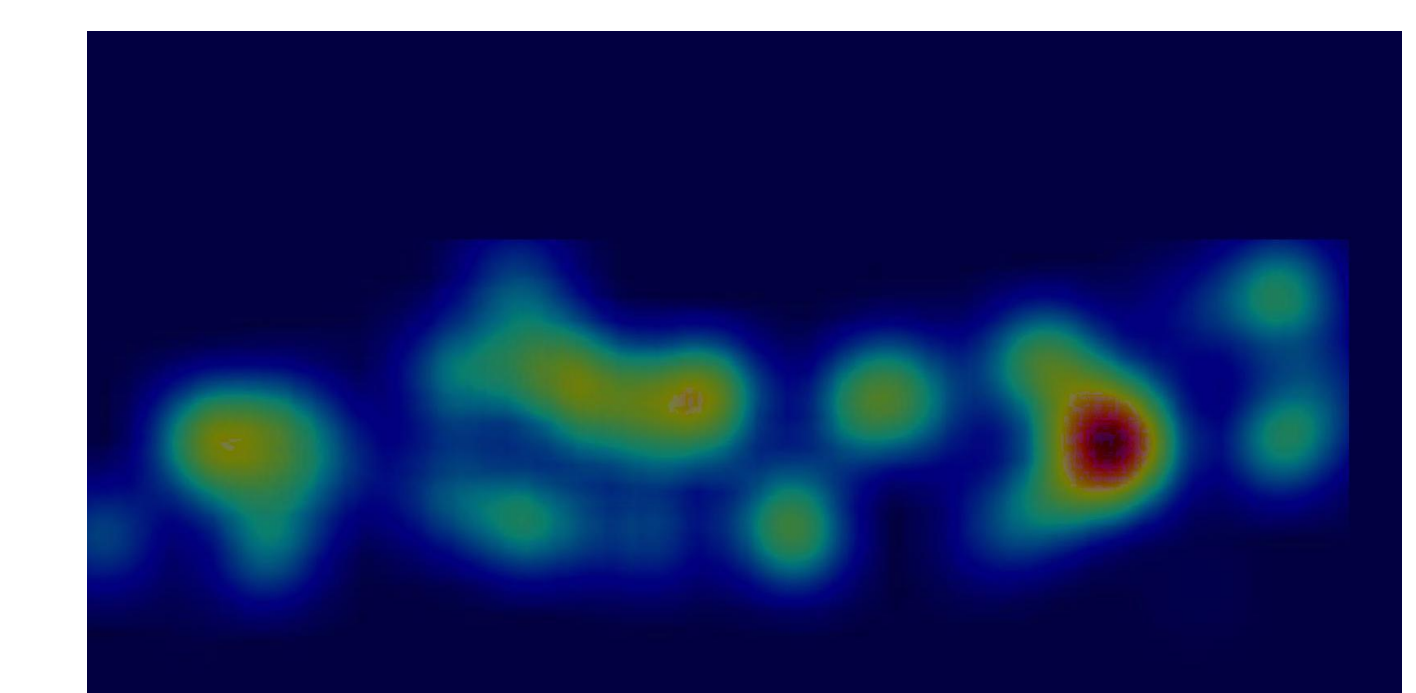
Predicted Output



Ground Truth



Saliency Map



Gaze Plot

- Saliency map showing the pixel-wise gradient of each pixel in the input image
- Illustrates the relative contribution of each pixel

Conclusion

- Multi-headed Approach works best
- Most large dental datasets aren't public
- Future Research
 - Exploring textual descriptions through Natural Language Processing
 - Vision Transformers
 - Self-Supervised Learning