**Analysis:** The probabilistic maps obtained were first thresholded at 0.5 from a range of [0,1]. Then the remaining connected components (CC) where analyzed (Figures 2.c and 5). The number of CC per tooth varied from a minimum number of 315 voxels to a maximum of 1000 voxels. Cracked teeth have a broader spread of CC sizes (between 315 and 1000 voxels) than healthy teeth (between 315 and 800 voxels). In this image, cracks tend to show up as long, thin connected components. Images that have no cracks have fewer, smaller connected components relative to those that have cracks present. In order to detect cracks in teeth, we chose distance-weighted discrimination (DWD) as our detection method and validated the results with direction-projection-permutation (DiProPerm) hypothesis tests. DWD was developed as an improved binary linear classifier to support vector machines (SVM), which finds a hyperplane that better separates the datasets (that is especially good for datasets with high heterogeneous?)1. DiProPerm2 is a permutation-based hypothesis test that assesses the chance that the observed degree of separation happened as a result of expected random variation. It was developed with DWD in mind as an area of application, but it represents a general framework of nonparametric hypothesis testing built to discern typical and atypical behavior in high-dimensional settings. Figure 5.a) and b) shows the results of classifying the CC elements found in the 14 healthy (black dots) and 31 cracked (red dots) teeth imaged with microCT images, illustrated by projecting the each teeth in the two principal component directions of the data, and with the multidimensional hyperplane illustrated in light green (Figure 5.a). Probability distribution functions of the data projected into the DWD separating hyperplane show no clear separation between healthy and cracked teeth. Thus, DiProPerm did not result in significant separation results (Figure 5.b, *p*=0.74). In hr-CBCT (not shown?), we find that despite having elevated connected component numbers there are many more false positives and thus much more overlap between groups, indicating even more limited detection ability.

Chart

Description automatically generated

**Figure 5. Phase I results. DWD classification of** connected component analysis obtained from our machine learning algorithm for a) microCT and for c) synthetically generated hr-CBCT. Probability distribution functions of data projections into the DWD separating hyperplane and permutation tests using DiProPerm for b) microCT and for d) synthetically generated hr-CBCT.

Detailed analysis into the acquisition parameters of the hr-CBCT revealed that the endodontic mode often uses smoothing filters intended to provide visually appealing images without the presence of artifacts. These filters are useful to better perceive the gross tooth structure and generate 3D reconstructions but can destroy fine features like microfractures as illustrated by Figure 6. To validate this finding, we tested our algorithm on a previously published dataset of synthetically generated cracks in hr-CBCT12.

Figure 5.c) and d) shows the results of classifying the CC elements found in the 19 synthetically cracked (red dots) and 6 healthy (blue dots) hr-CBCT images, showing appropriate separation of CC obtained from healthy and cracked teeth using a similar plane projection. DiProPerm testing showed significant separation of the two groups (Figure 5.d. *p*=0.026) even for this small sample size. Results with this data show that our approach can be applied to hr-CBCT (clinically) when the images are not over-processed. This work has now been accepted for podium presentation at the Society of Photo-optical Instrumentation Engineers in Medical Imaging Conference (SPIE MI 2021), one of the most prestigious in our field.

We now plan to collect images without additional filters moving forward. As noted in Figure 6, it is hard to distinguish between the crack and the streak artifact. The goal of specific Aim 1 is to improve the technology to be robust to radiopaque materials such as restorations and root canal obturants (filling materials). In specific Aim 3 we will validate our technology clinically by including teeth with full coverage restorations and root canal obturants.

1 Marron, J. S., Todd, M. J. & Ahn, J. Distance-Weighted Discrimination. *Journal of the American Statistical Association* **102**, 1267-1271, doi:10.1198/016214507000001120 (2007).

2 Wei, S., Lee, C., Wichers, L. & Marron, J. S. Direction-Projection-Permutation for High-Dimensional Hypothesis Tests. *Journal of Computational and Graphical Statistics* **25**, 549-569, doi:10.1080/10618600.2015.1027773 (2016).