Датасет состоит из архива изображений ротовой полости и разметки в формате xml.

Разметка состоит из двух классов: teeth и caries.

Фотографии могут быть размечены двумя способами.

Первый способ, когда разметка caries находится внутри разметки teeth и определяет границы самого кариеса.   
Изображение выглядит как текст

Автоматически созданное описание

Второй способ, когда разметкой выделяется сам зуб с кариесом.

Изображение выглядит как текст, легкий

Автоматически созданное описание

В датасете они не разделены, но есть определенная закономерность. При первом варианте выделение кариеса находится внутри выделения зуба, а во втором выделение зуба находится внутри выделения зуба с кариесом.