

Important notes:

Do **NOT** write outside the grey boxes. Any text or images outside the boxes **will** be deleted.

Do **NOT** alter the structure of this form. Simply enter your information into the boxes. The form will be automatically processed – if you alter its structure your submission will not be processed correctly.

Do not include keywords – you can add them when you submit the abstract online.

Title:

From bolus formation analysis to scalable segmentation: Integrating pseudo-labeling with Mask R-CNN for chewing studies

Authors & affiliations:

<i>Shiyi Ren</i>	<i>The University of Auckland</i>
<i>Bangxiang Chen</i>	<i>Xi'an Jiaotong-Liverpool University</i>
<i>Jaspreet Dhupia</i>	<i>The University of Auckland</i>
<i>Martin Stommel</i>	<i>Auckland University of Technology</i>
<i>Peter Xu</i>	<i>The University of Auckland</i>

Abstract: (Your abstract must use **Normal style** and must fit in this box. Your abstract should be no longer than 300 words. The box will 'expand' over 2 pages as you add text into it.)

Preparation of Your Abstract

1. The title should be as brief as possible but long enough to indicate clearly the nature of the study. Capitalise the first letter of the first word **ONLY** (place names excluded). No full stop at the end.
2. Abstracts should state briefly and clearly the purpose, methods, results and conclusions of the work.

Introduction: Clearly state the purpose of the abstract

Methods: Describe your selection of observations or experimental subjects clearly

Results: Present your results in a logical sequence

Discussion: Emphasize new and important aspects of the study and conclusions that are drawn from them

Chewing is essential for digestion, fragmenting food and forming a cohesive bolus suitable for swallowing. In our earlier work, we introduced a non-destructive, image-based approach using Mask Region-based Convolutional Neural Networks (Mask R-CNN) to analyze bolus formation dynamics. By benchmarking backbone architectures, we showed that ResNet101 achieved the most reliable segmentation of bolus particles, improving upon traditional particle size distribution methods and establishing a computational foundation for non-invasive chewing analysis.

In this study, we extend the framework to address a major bottleneck: limited annotated training data. We combined manually labelled key frames with pseudo-labelled intermediate frames generated via temporal propagation, thereby expanding dataset size without proportional manual effort. This hybrid approach allowed the model to learn from both accurate annotations and temporally dense pseudo labels. Training was performed with early stopping to prevent overfitting, and evaluation was conducted on an independent manually annotated validation set.

Results showed that pseudo-labelled augmentation improved overall recall and temporal robustness yet introduced noise that reduced precision for minority classes. Compared with the baseline trained solely on manual labels, the augmented model better captured large-scale bolus transitions but remained challenged by fine-grained minority structures. These findings highlight both the potential and the pitfalls of pseudo-labelling in biomedical and food imaging applications.

This work creates a narrative from methodological innovation in bolus segmentation to practical strategies for scalable data preparation. It lays the groundwork for real-time segmentation systems in robotic chewing simulators and personalized food design, advancing the application of deep learning in food oral processing research.