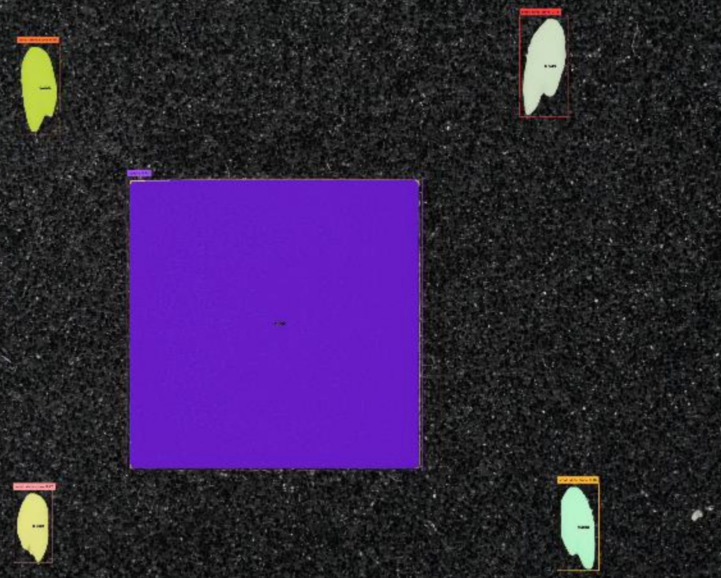
PICES-2024 Annual Meeting, Honolulu, USA

Automatic detection and measurement of otolith using zero-shot learning



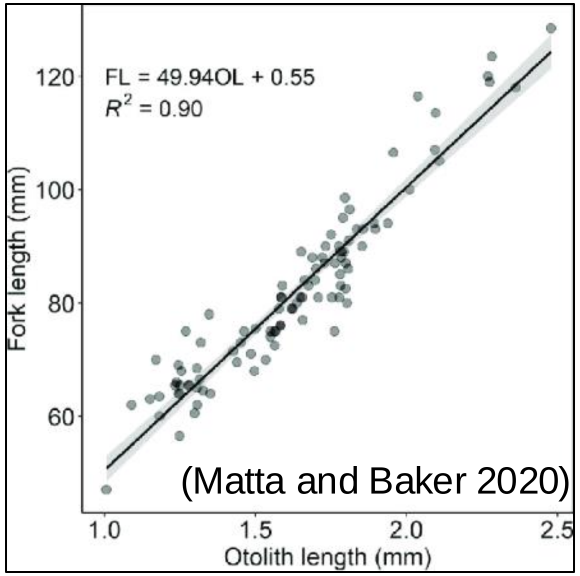
Masahiro **Manano**[1], Yuko Hiraoka[1], Yasutoki Shibata[1], Yuka Iwahara[1], Daisaku Masuda[2]

[1]Japan Fisheries Research and Education Agency, Japan

[2]Nagasaki Prefectural Institute of Fisheries, Japan

Introduction

- **Prey size composition of piscivorous fish**
  - Basic ecological data of predators[1,2]
  - It can also help us understand the status of prey stocks[1].
  - However, measuring the length of prey in the stomach contents is often difficult due to digestion.



(Matta and Baker 2020)

The body length of prey can be predicted from the length of the otoliths[3,4].

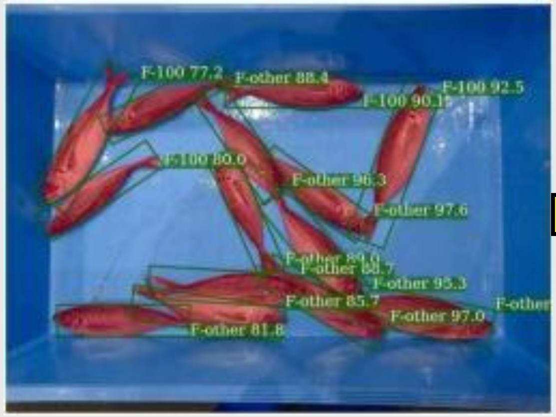
[1]Beveren et al. 2017 [2]Ménard et al. 2007 [3,4]Battaglia et al. 2010, 2015

Introduction

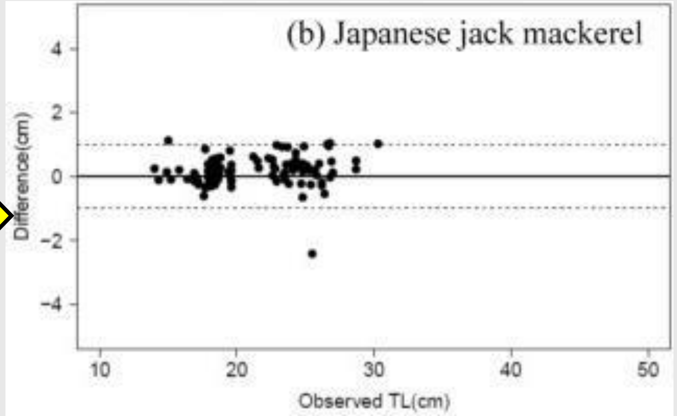
- **Challenges with length measurement**
  - Time-consuming.
  - Measurement errors and biases between different measurers.
  - Measurement errors and biases within the same measurer over time.

Introduction

- **Challenges with length measurement**
  - Time-consuming.
  - Measurement errors and biases between different measurers.
  - Measurement errors and biases within the same measurer over time.

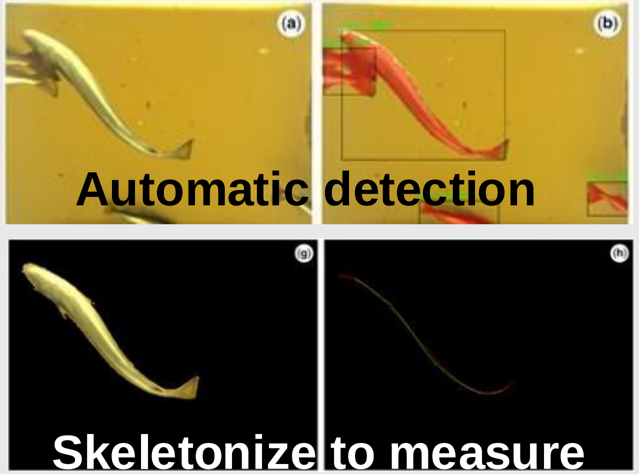


Automatic detection



Length prediction

(Adapted from Shibata et al. 2024)



(Adapted from Garcia et al. 2020)

The application of deep learning technology, which aims to reduce labor and improve reproducibility, is being explored across various fields[5,6,7,8].

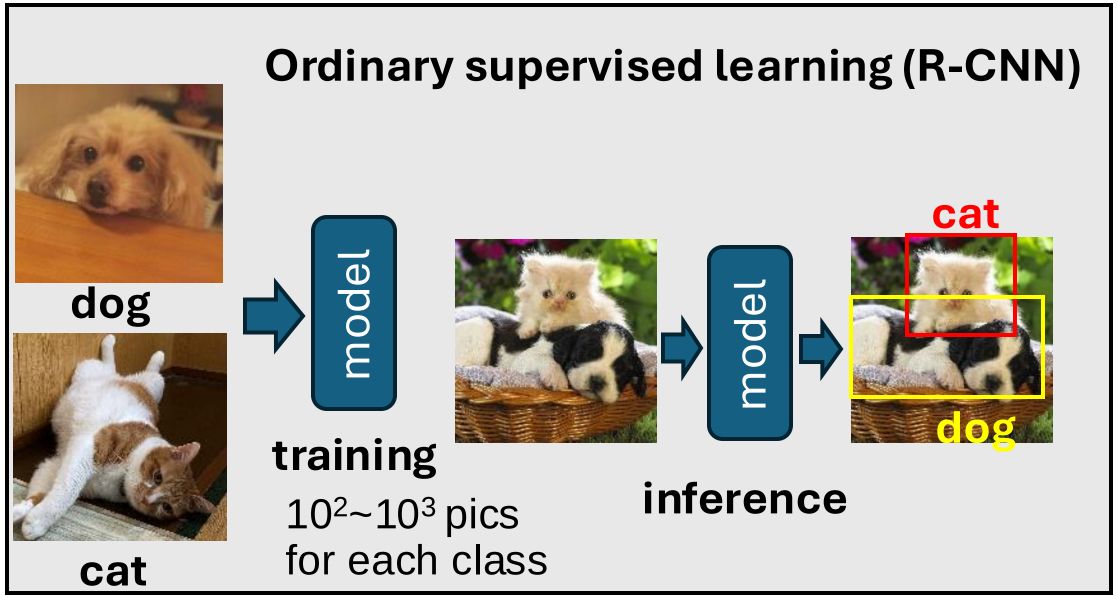
[5]Salman et al. 2019 [6]Moen et al. 2018 [7]Palmer et al. 2021 [8]Tseng and Kuo 2020

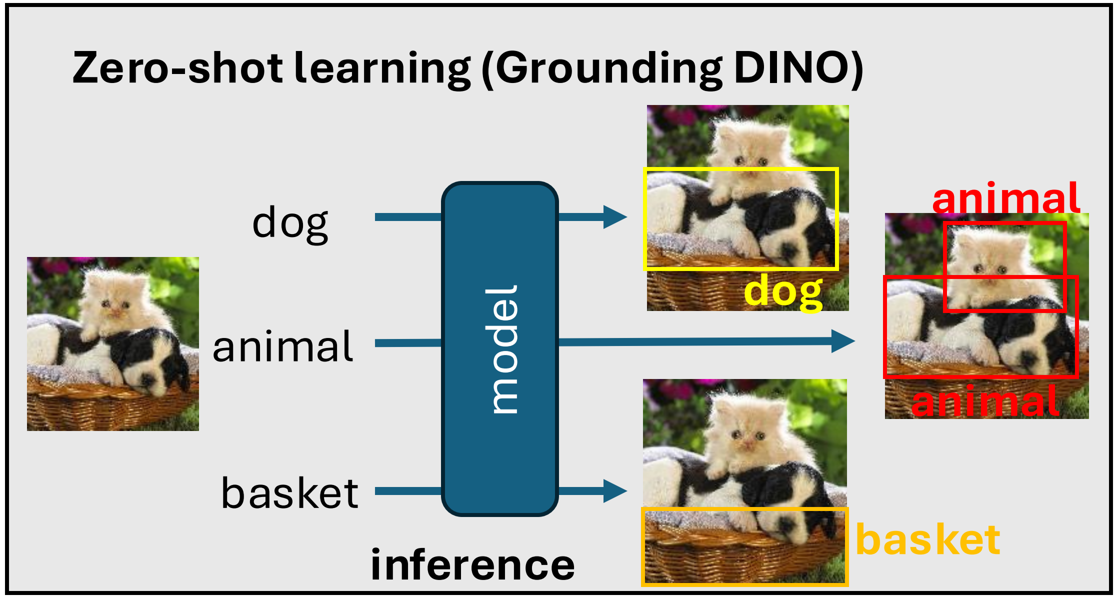
Introduction

- The process of collecting and preparing data for supervised machine learning requires significant time, effort, and cost.

⇨ Foundation model

- Models trained on broad data that can be adapted to a wide range of downstream tasks[9]
- It enables **zero-shot learning**





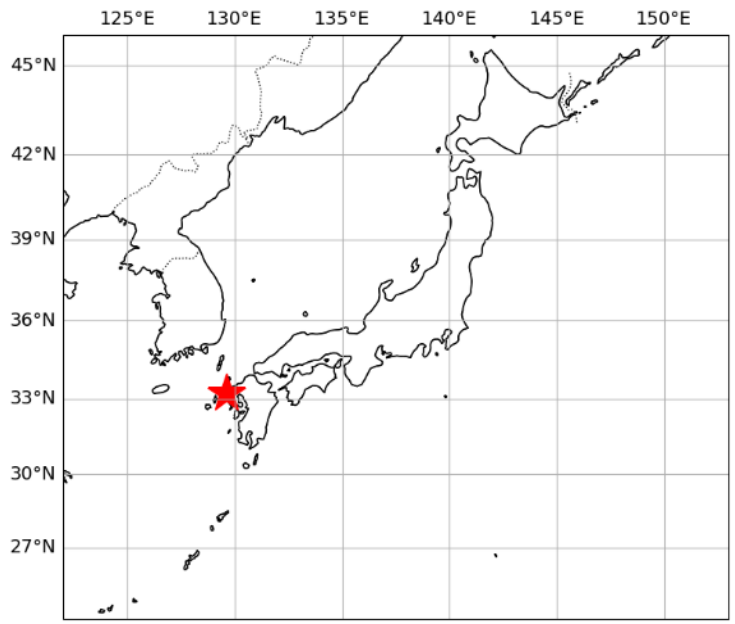
[9]Bommasani et al. 2022

Objective

This study aims to develop a model that automatically detects multiple otoliths in images and predict their lengths using zero-shot learning.

- This is expected to significantly reduce the effort and time required for otolith length measurement while also improving the reproducibility of the measurements.

Materials and Methods



Sampling

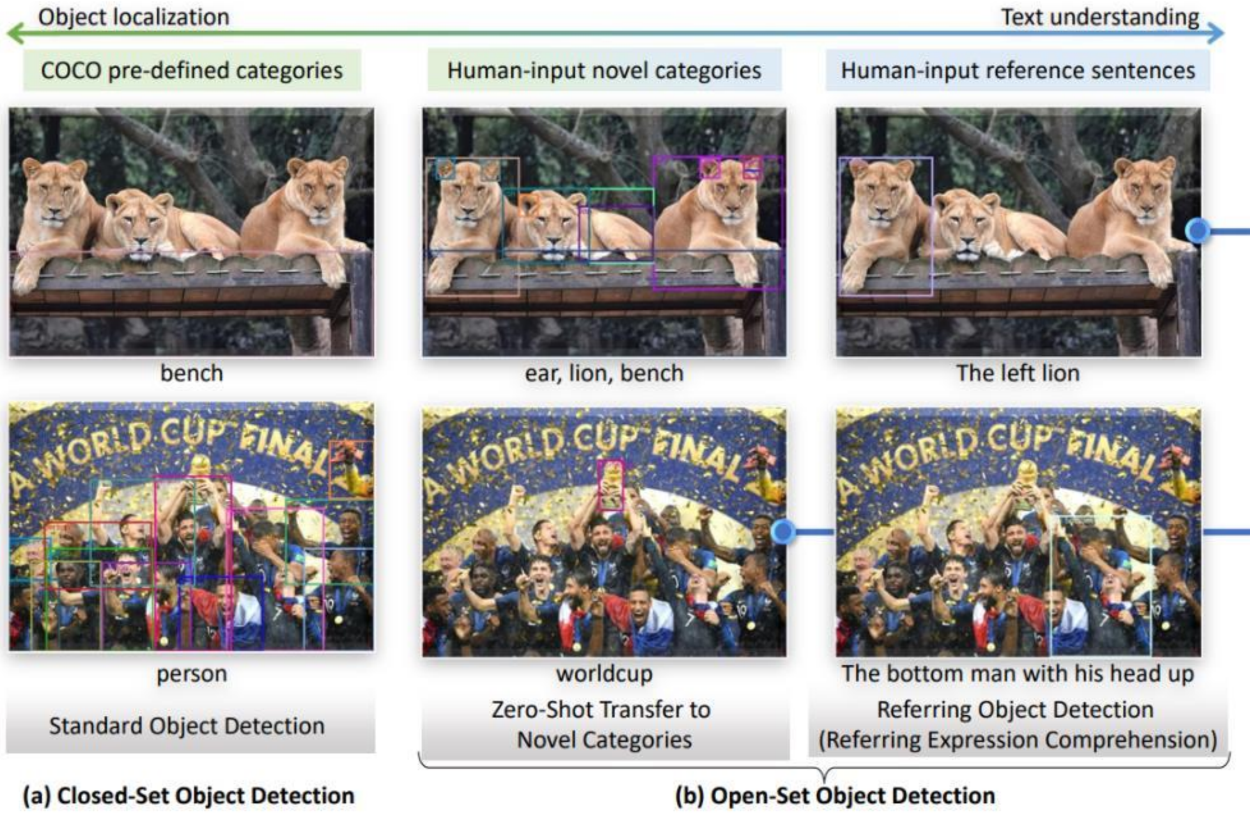
- Sasebo city, Nagasaki prefecture, Japan
- Japanese anchovy (*Engraulis japonica*)

Date	n	SL* (mean±SD; mm)	OL* (mean±SD; mm)
2021/7/10	32	67.59 ± 3.30	2.140 ± 0.802
2021/11/29	44	42.99 ± 6.24	2.278 ± 0.179
2022/4/19	45	99.82 ± 12.11	1.204 ± 0.305
Total	**121**	**70.63 ± 25.94**	**1.959 ± 0.254**

* SL : scaled length, OL : otolith length

Materials and Methods

Grounding DINO[10]



(Adapted from Liu et al. 2023)

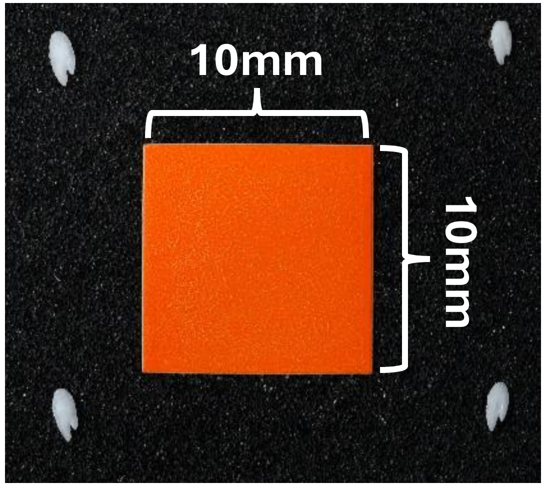
Segment Anything[11]

(Adapted from Kirillov et al. 2023)

Object detection and segmentation based on text prompt <u>with no training</u>

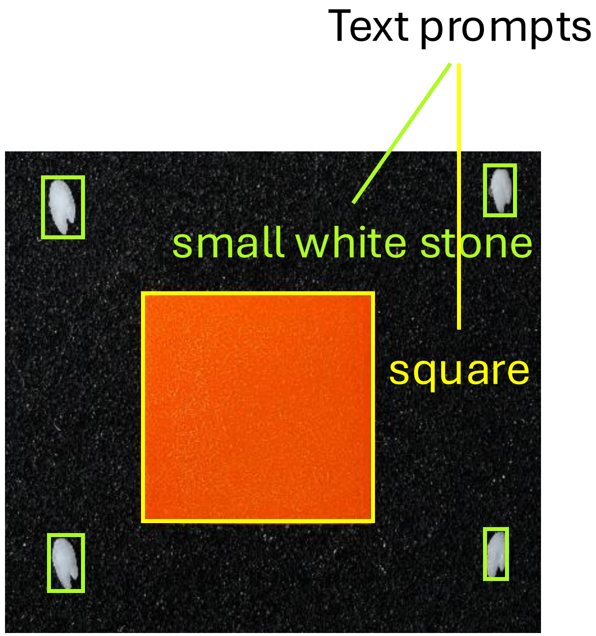
[10]Liu et al. 2023 [11]Kirillov et al. 2023

Materials and Methods

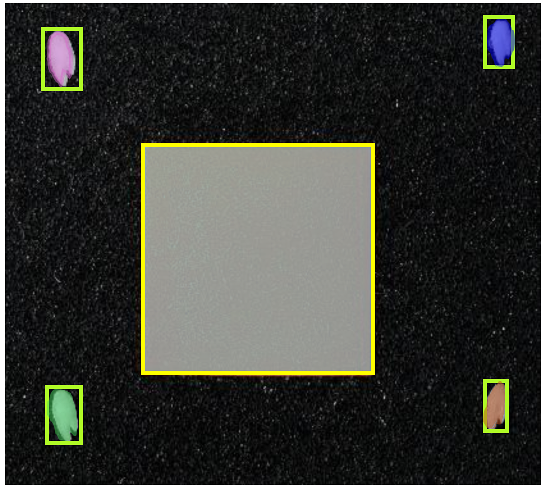


- Four otoliths in each images
- Black NR sponge

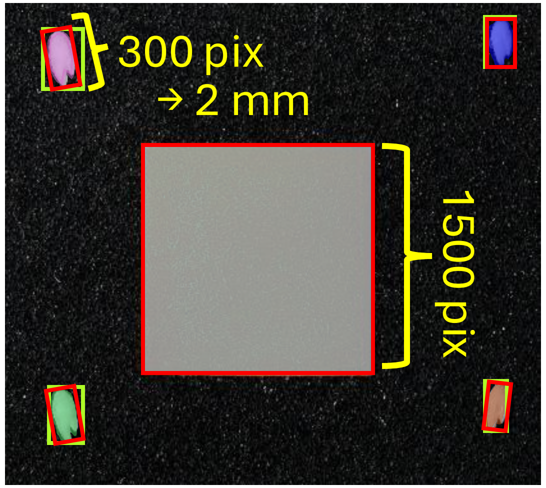
Grounding DINO
➤ **Detection**



Segment Anything
➤ **Segmentation**



◆ **Minimum bounding box**
◆ **Counting pixels of the long side**



Output
csv file

Materials and Methods

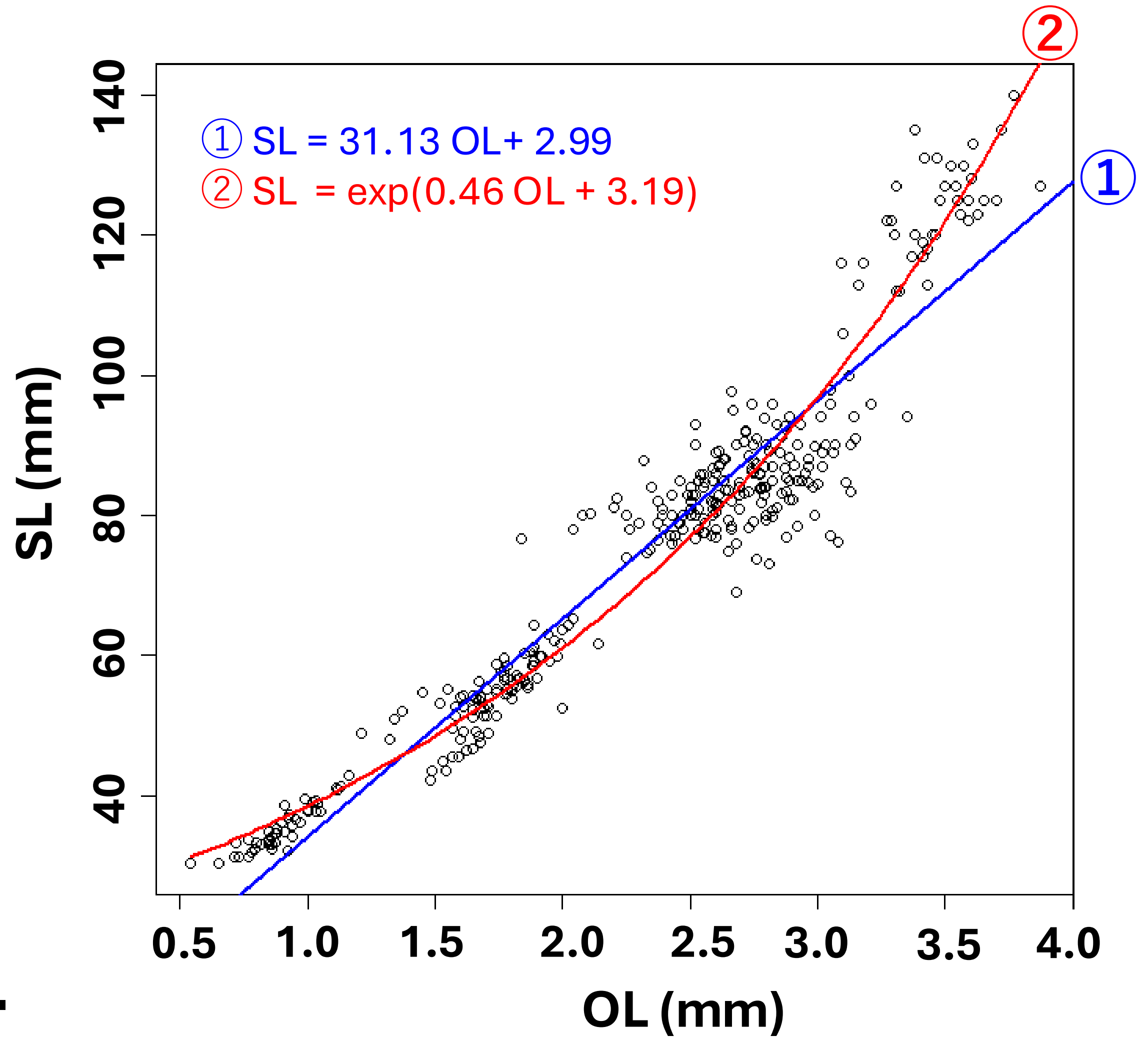
OL-SL relationship
to convert predicted OLs to SLs

Sampling site	Nagasaki pref.
Year	2013, 2017
N	365
SL (mm)	30.36 - 140.00
OL (mm)	0.54 – 3.87

① Linear model (R^2 = 0.91)

② Log model (R^2 = 0.95)

Model② was adopted in this study.



Materials and Methods

Data analysis on predicted OLs (similarly for SLs)

- The prediction accuracy of the OLs using deep learning models was evaluated by **relative bias (RB)** and **RMSE**.

1. Relative Bias (RB)

$$RD_i = \frac{\overline{OL}_i - OL_i}{OL_i}$$

$$i = 1, 2, 3, \ldots, n$$

$$RB = \frac{1}{n}\sum RD_i$$

$$i = 1, 2, 3, \ldots, n$$

2. Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum(\overline{OL}_i - OL_i)^2}$$

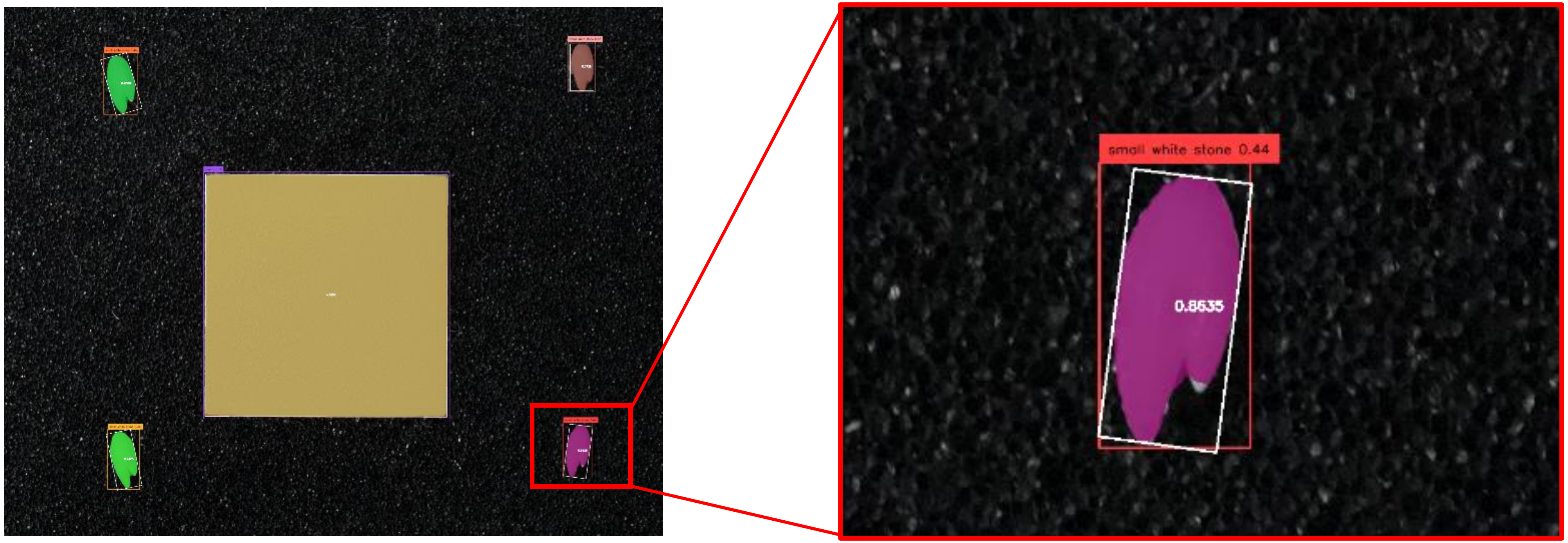
$$i = 1, 2, 3, \ldots, n$$

OL: observed otolith length, $\overline{OL}$: predicted otolith length, $n = 121$

Results

Detection and segmentation

- **All 121 otoliths were automatically detected and segmented** using Grounding DINO and Segment Anything **without any training**, **with no over- or underdetection**.

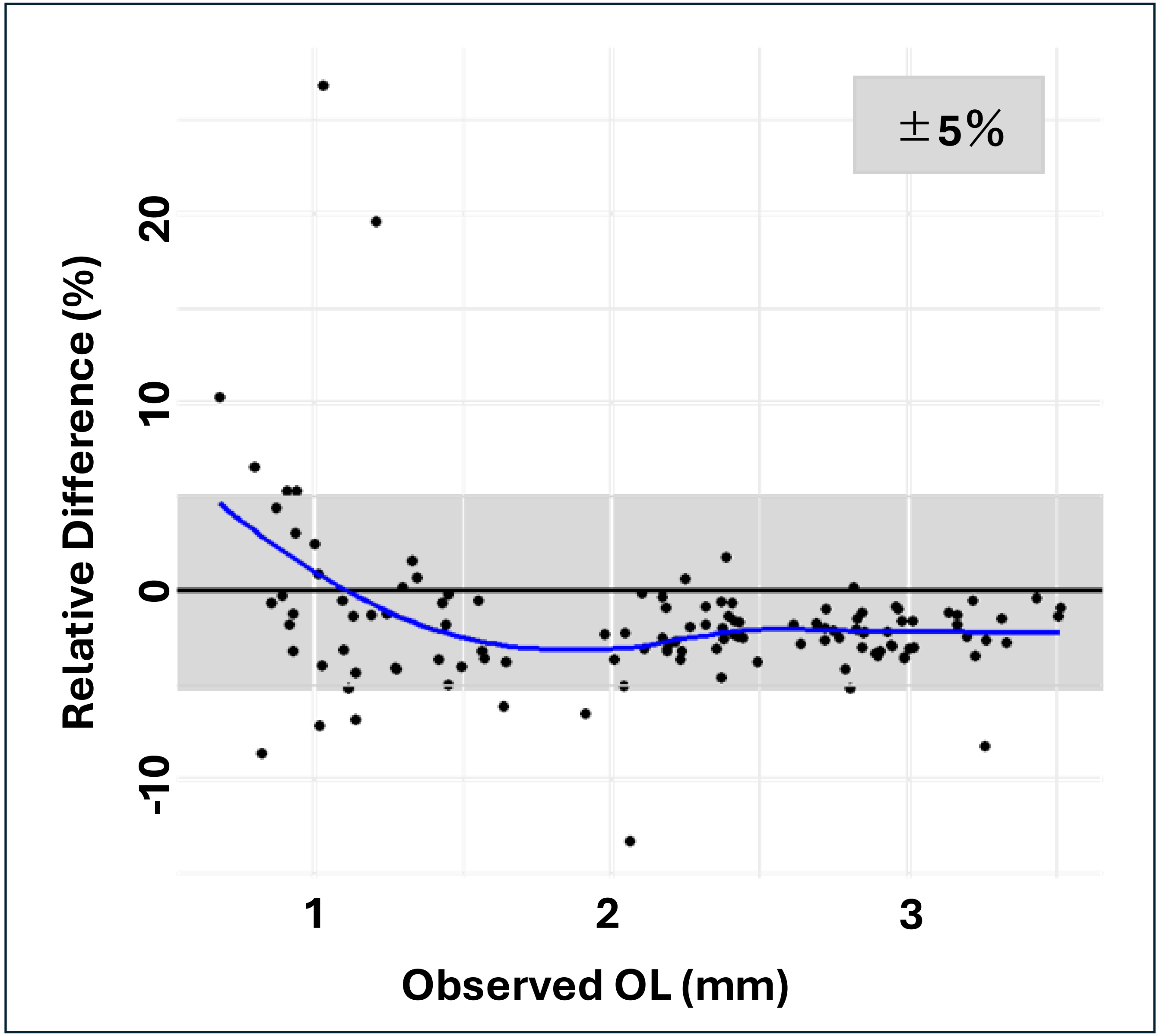


Results | OL

Relative Bias and RMSE

- **RB = -1.59 %**
- **RMSE = 0.08 (mm)**

The otolith length of **103 out of 121 individuals were** predicted with an error of **less than 5%** compared to the actual measurements.



The blue line represents the trend in the plot using LOESS.

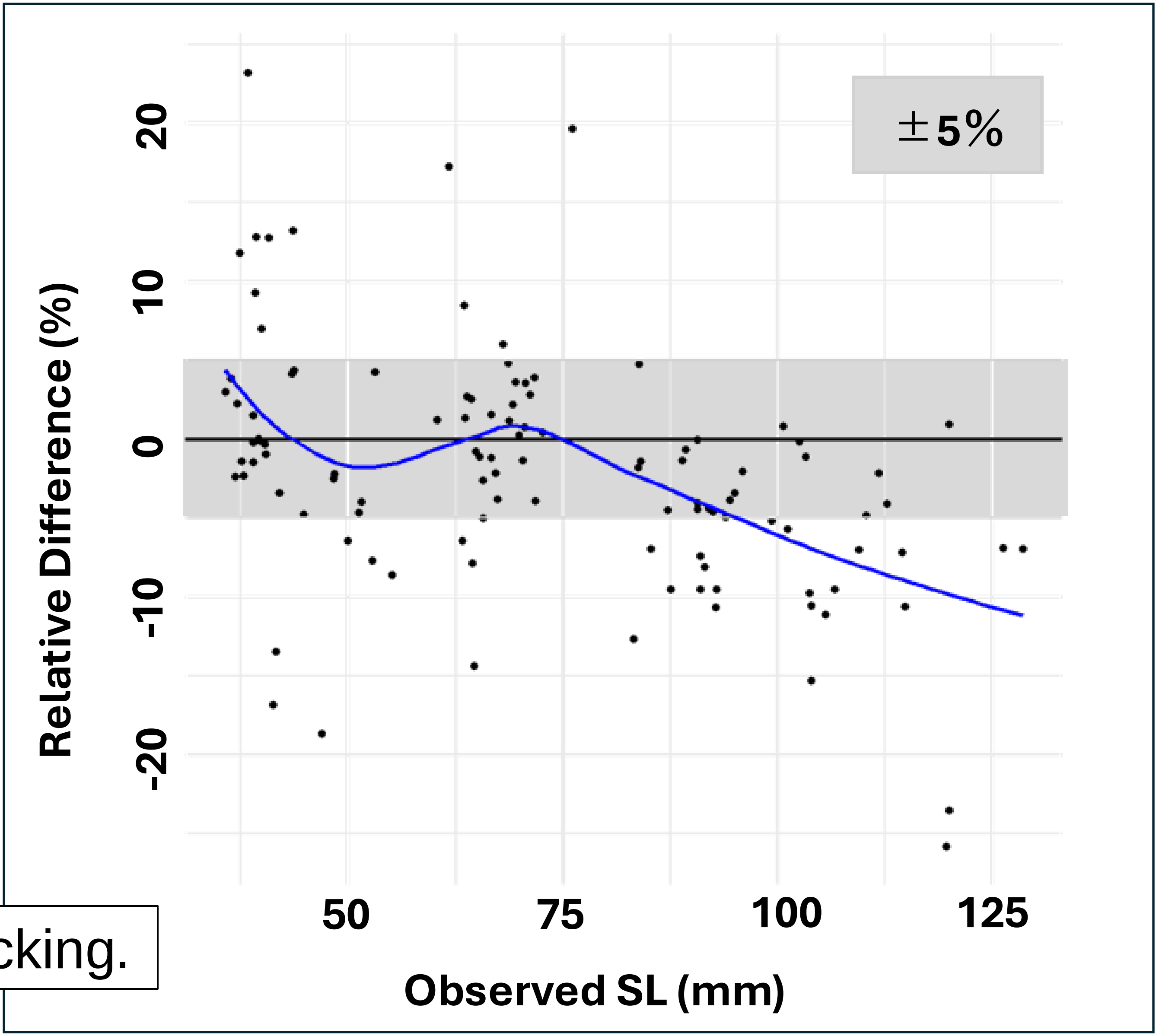
Results | SL

Relative Bias and RMSE

- **RB = -2.09 %**
- **RMSE = 6.55 (mm)**

The scaled length of **69 out of 111 individuals** were predicted with an error of **less than 5%** compared to the actual measurements.

Ten otoliths were excluded due to cracking.

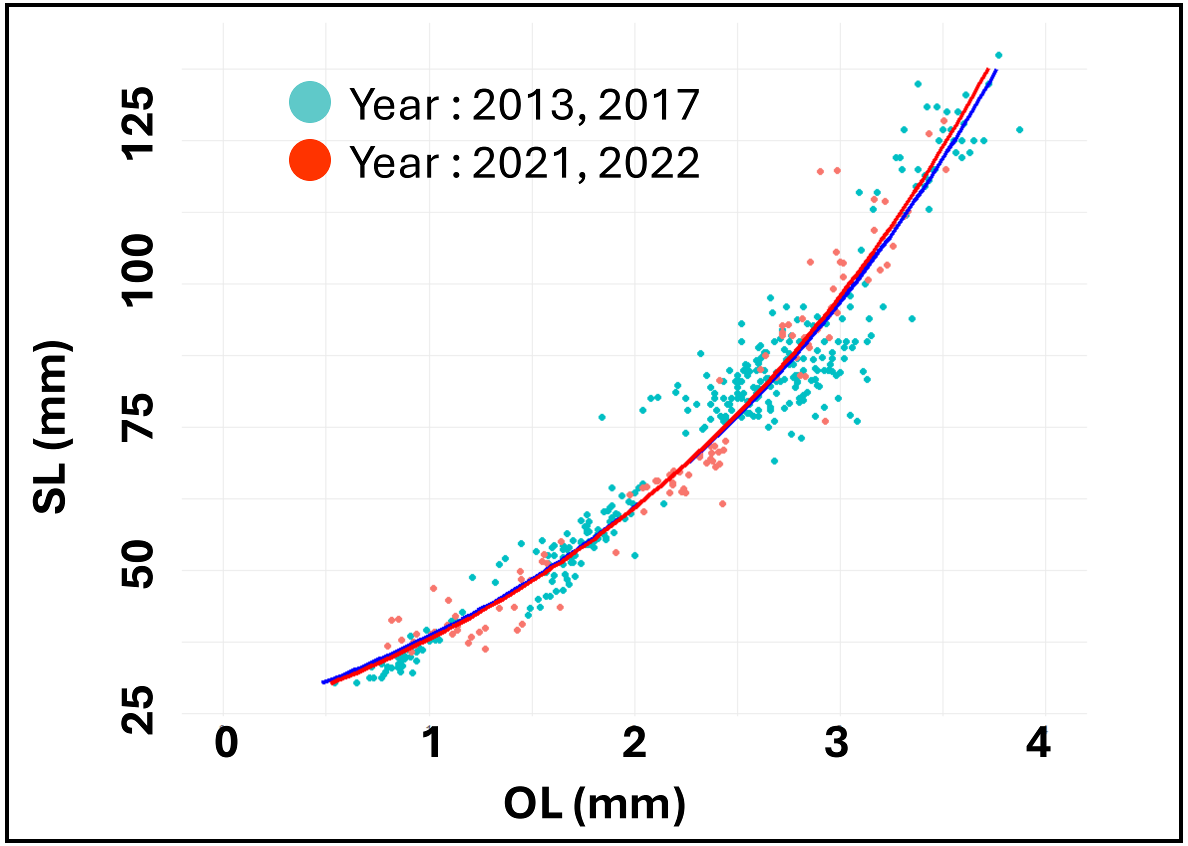


The blue line represents the trend in the plot using LOESS.

Discussion

- **The predicted OLs were accurate.**
- The RMSE of the predicted SLs was higher than that of the predicted OLs.
  - ➢ This was likely caused by the variance in the OL-SL relationship.
- The RD in the predicted SLs decreased as the observed SLs increase.
  - ➢ There was a difference in the OL-SL relationship between the two sampling situations.

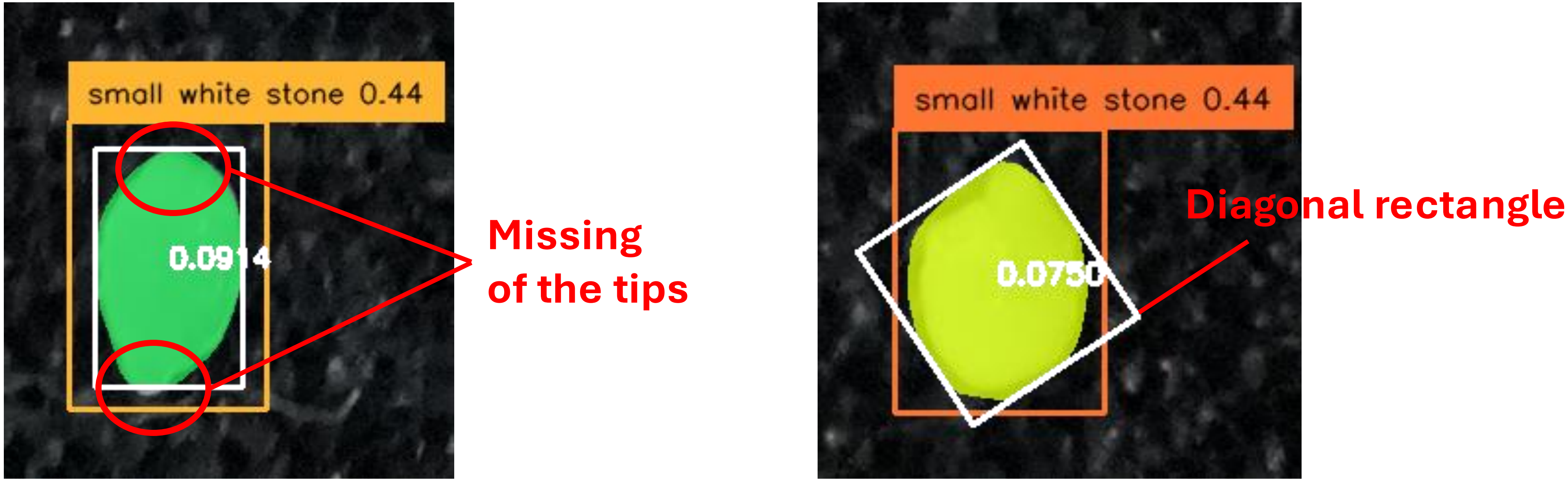
	RB (%)	RMSE (mm)
OL	-1.59	0.10 (3.7% of mean OLs)
SL	-2.09	6.55 (9.1% of mean SLs)



Discussion

Masks generated by SAM appeared accurate, but...

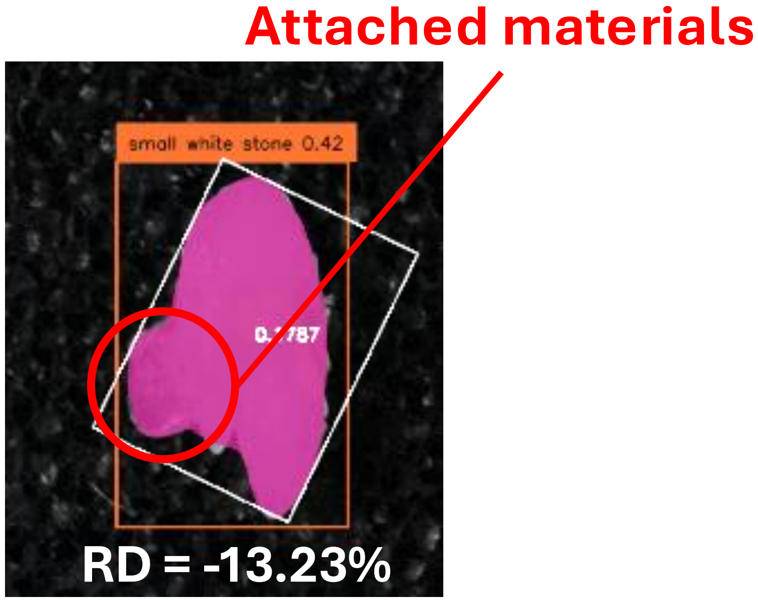
- The predicted OLs were slightly negatively biased.
  - ➢SAM sometimes fails to create a mask for the tip of the otolith.
  - ➢The long side of the minimum bounding rectangle did not exactly align with the OL.

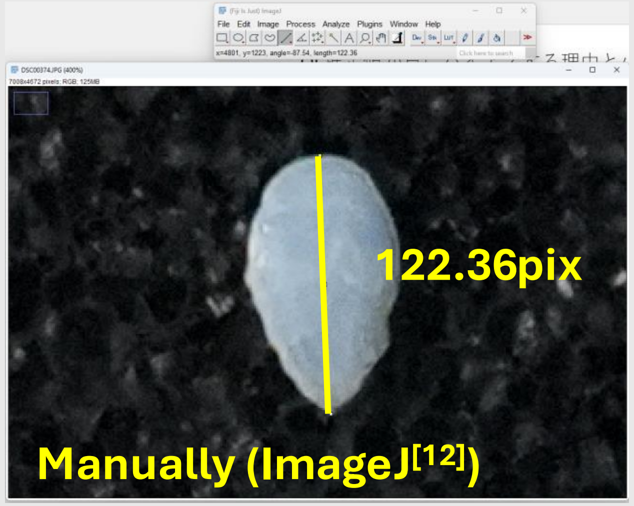


Discussion

Masks generated by SAM appeared accurate, but...

- The predicted OLs for smaller otoliths had larger variance.
  - ➢ Masks that includes attached materials
  - ➢ Observation error
  - ➢ The smaller the otolith, the greater the impact of a 1-pixel shift on the RB.





	pix	mm	RB
OL	-	0.797	-
$\overline{OL}$	120	0.850	6.6%
Manually measured	122.36	0.867	8.8%

[12] ImageJ (Version 1.54f). National Institutes of Health, USA

Conclusion

- **We have developed a model for automatic otolith detection and length prediction without the need for training.**
  - ➢This serves as the foundation for reconstructing prey body length from hard tissues in the stomachs of piscivorous fish.
  - ➢This technology is expected to significantly reduce the time and effort required for measurement.
  - ➢Further experiments are needed to assess its reproducibility.