

## Supporting Information

**Table S1.** Post hoc comparisons for the interaction between accuracy metric and image source (validation set).

Comparison	<i>t</i> (188)	<i>p</i> <sub>adj.</sub>	Cohen's <i>d</i>
<b>FreqIoU, iNat &gt; FreqIoU, Randall</b>	<b>14.839</b>	< .001	<b>1.077</b>
<b>FreqIoU, iNat &gt; MeanAcc, iNat</b>	<b>4.035</b>	<b>0.002</b>	<b>0.293</b>
FreqIoU, iNat - MeanAcc, Randall	2.318	0.590	0.168
<b>FreqIoU, iNat &gt; MeanIoU, iNat</b>	<b>16.331</b>	< .001	<b>1.185</b>
<b>FreqIoU, iNat &gt; MeanIoU, Randall</b>	<b>17.004</b>	< .001	<b>1.234</b>
<b>FreqIoU, iNat &lt; PxAcc, iNat</b>	<b>-11.407</b>	< .001	<b>-0.828</b>
FreqIoU, iNat - PxAcc, Randall	-0.771	1.000	-0.056
<b>FreqIoU, Randall &lt; MeanAcc, iNat</b>	<b>-11.836</b>	< .001	<b>-0.859</b>
<b>FreqIoU, Randall &lt; MeanAcc, Randall</b>	<b>-32.078</b>	< .001	<b>-2.327</b>
FreqIoU, Randall - MeanIoU, iNat	-2.687	0.212	-0.195

<b>FreqIoU, Randall &gt; MeanIoU, Randall</b>	<b>5.547</b>	< .001	<b>0.402</b>
<b>FreqIoU, Randall &lt; PxAcc, iNat</b>	<b>-23.327</b>	< .001	<b>-1.692</b>
<b>FreqIoU, Randall &lt; PxAcc, Randall</b>	<b>-39.993</b>	< .001	<b>-2.901</b>
MeanAcc, iNat - MeanAcc, Randall	-0.685	1.000	-0.050
<b>MeanAcc, iNat &gt; MeanIoU, iNat</b>	<b>12.296</b>	< .001	<b>0.892</b>
<b>MeanAcc, iNat &gt; MeanIoU, Randall</b>	<b>14.001</b>	< .001	<b>1.016</b>
<b>MeanAcc, iNat &lt; PxAcc, iNat</b>	<b>-15.442</b>	< .001	<b>-1.120</b>
<b>MeanAcc, iNat &lt; PxAcc, Randall</b>	<b>-3.774</b>	<b>0.005</b>	<b>-0.274</b>
<b>MeanAcc, Randall &gt; MeanIoU, iNat</b>	<b>9.834</b>	< .001	<b>0.713</b>
<b>MeanAcc, Randall &gt; MeanIoU, Randall</b>	<b>37.625</b>	< .001	<b>2.730</b>
<b>MeanAcc, Randall &lt; PxAcc, iNat</b>	<b>-10.806</b>	< .001	<b>-0.784</b>
<b>MeanAcc, Randall &lt; PxAcc, Randall</b>	<b>-7.914</b>	< .001	<b>-0.574</b>
<b>MeanIoU, iNat &gt; MeanIoU, Randall</b>	<b>4.852</b>	< .001	<b>0.352</b>

<b>MeanIoU, iNat &lt; PxAcc, iNat</b>	<b>-27.739</b>	<b>&lt; .001</b>	<b>-2.012</b>
<b>MeanIoU, iNat &lt; PxAcc, Randall</b>	<b>-12.923</b>	<b>&lt; .001</b>	<b>-0.938</b>
<b>MeanIoU, Randall &lt; PxAcc, iNat</b>	<b>-25.492</b>	<b>&lt; .001</b>	<b>-1.849</b>
<b>MeanIoU, Randall &lt; PxAcc, Randall</b>	<b>-45.540</b>	<b>&lt; .001</b>	<b>-3.304</b>
<b>PxAcc, iNat &gt; PxAcc, Randall</b>	<b>7.717</b>	<b>&lt; .001</b>	<b>0.560</b>

*Note.* Student's *t*-test. Bonferroni-corrected (for multiple comparisons) post hoc paired-samples *t*-tests comparing the four image segmentation validation metrics; pixel accuracy (PxAcc, Eq. 1), mean accuracy (MeanAcc, Eq. 2), mean intersection over union (MeanIoU, Eq. 3), and frequency weighted intersection over union (FreqIoU, Eq. 4), from either iNaturalist (iNat) or J.E. Randall's fish photos collection (Randall). Cohen's *d* does not correct for multiple comparisons. Significant comparisons are bolded

### ***Quantitative Image Segmentation Evaluation Metrics for Novel Test Set***

We constructed a test dataset of 60 novel images (i.e., all images were new images not included in the original training/validation datasets): 30 images were J.E. Randall's fish images in left-lateral view on a uniform black backdrop, and 30 images were naturalistic images of fishes in various views on complex noisy backdrops. All 60 novel test images had mean IoU scores greater than 50% ( $M = 94.7\%$ ,  $SD = 1.1\%$ , minimum = 91.8%, maximum = 96.9%), indicating excellent model-predicted segmentation masks compared to manually drawn reference masks. Comparing across image segmentation metrics and image sources, we found a significant main effect of evaluation metric on accuracy,  $F(1.42, 82.60) = 344.65$ ,  $p_{adj.} < .001$ , suggesting that independent of image source (iNaturalist, J.E. Randall), accuracy metrics varied significantly from one another. We also found a significant main effect of image source,  $F(1, 58) = 65.19$ ,  $p < .001$ , such that regardless of accuracy metric, images from iNaturalist were generally segmented with higher accuracy ( $M = 96.9\%$ ,  $SD = .7\%$ ) than were J.E. Randall's images ( $M = 95.5\%$ ,  $SD = .5\%$ ),  $t(58) = 8.07$ , Cohen's  $d = 1.04$ ,  $p_{adj.} < .001$ . Lastly, we uncovered a significant interaction between accuracy metric and image source,  $F(1.42, 82.60) = 55.38$ ,  $p_{adj.} < .001$ . Bonferroni-corrected post hoc paired-samples  $t$ -tests for the significant main effect of metric revealed frequency weighted IoU to be significantly less than pixel and mean accuracy, but significantly higher than mean IoU. Additionally, pixel accuracy was significantly higher than both mean accuracy and mean IoU, and mean accuracy was significantly higher than mean IoU (Table S2).

**Table S2.** Post hoc comparisons for the main effect of evaluation metric (test set).

Comparison	<i>t</i> (58)	<i>p</i> <sub>adj.</sub>	Cohen's <i>d</i>
FreqIoU < MeanAcc	-3.58	.004	-.46
FreqIoU > MeanIoU	7.66	< .001	.99
FreqIoU < PxAcc	-19.91	< .001	-2.57
MeanAcc > MeanIoU	18.19	< .001	2.35
MeanAcc < PxAcc	-9.38	< .001	-1.21
MeanIoU < PxAcc	-35.05	< .001	-4.53

*Note.* Student's *t*-test. Bonferroni-corrected (for multiple comparisons) post hoc paired-samples *t*-tests comparing the four image segmentation validation metrics; pixel accuracy (PxAcc, Eq. 1), mean accuracy (MeanAcc, Eq. 2), mean intersection over union (MeanIoU, Eq. 3), and frequency weighted intersection over union (FreqIoU, Eq. 4). Cohen's *d* does not correct for multiple comparisons

**Table S3.** Post hoc comparisons for the interaction between accuracy metric and image source (test set).

Comparison	t(58)	p <sub>adj.</sub>	Cohen's d
<b>FreqIoU, iNat &gt; FreqIoU, Randall</b>	<b>12.809</b>	< .001	<b>1.654</b>
<b>FreqIoU, iNat &gt; MeanAcc, iNat</b>	<b>3.487</b>	<b>0.017</b>	<b>0.450</b>
FreqIoU, iNat - MeanAcc, Randall	2.762	0.185	0.357
<b>FreqIoU, iNat &gt; MeanIoU, iNat</b>	<b>12.296</b>	< .001	<b>1.587</b>
<b>FreqIoU, iNat &gt; MeanIoU, Randall</b>	<b>14.597</b>	< .001	<b>1.884</b>
<b>FreqIoU, iNat &lt; PxAcc, iNat</b>	<b>-10.040</b>	< .001	<b>-1.296</b>
FreqIoU, iNat - PxAcc, Randall	-0.192	1.000	-0.025
<b>FreqIoU, Randall &lt; MeanAcc, iNat</b>	<b>-10.430</b>	< .001	<b>-1.347</b>
<b>FreqIoU, Randall &lt; MeanAcc, Randall</b>	<b>-14.727</b>	< .001	<b>-1.901</b>
<b>FreqIoU, Randall &lt; MeanIoU, iNat</b>	<b>-4.420</b>	< .001	<b>-0.571</b>
FreqIoU, Randall - MeanIoU, Randall	2.620	0.268	0.338
<b>FreqIoU, Randall &lt; PxAcc, iNat</b>	<b>-19.658</b>	< .001	<b>-2.538</b>
<b>FreqIoU, Randall &lt; PxAcc, Randall</b>	<b>-19.058</b>	< .001	<b>-2.460</b>
MeanAcc, iNat - MeanAcc, Randall	0.383	1.000	0.049
<b>MeanAcc, iNat &gt; MeanIoU, iNat</b>	<b>8.809</b>	< .001	<b>1.137</b>
<b>MeanAcc, iNat &gt; MeanIoU, Randall</b>	<b>12.218</b>	< .001	<b>1.577</b>
<b>MeanAcc, iNat &lt; PxAcc, iNat</b>	<b>-13.527</b>	< .001	<b>-1.746</b>
MeanAcc, iNat - PxAcc, Randall	-2.571	0.317	-0.332
<b>MeanAcc, Randall &gt; MeanIoU, iNat</b>	<b>5.626</b>	< .001	<b>0.726</b>

<b>MeanAcc, Randall &gt; MeanIoU, Randall</b>	<b>17.347</b>	< .001	2.239
<b>MeanAcc, Randall &lt; PxAcc, iNat</b>	<b>-9.612</b>	< .001	<b>-1.241</b>
<b>MeanAcc, Randall &lt; PxAcc, Randall</b>	<b>-4.331</b>	< .001	<b>-0.559</b>
<b>MeanIoU, iNat &gt; MeanIoU, Randall</b>	<b>6.208</b>	< .001	0.801
<b>MeanIoU, iNat &lt; PxAcc, iNat</b>	<b>-22.336</b>	< .001	<b>-2.884</b>
<b>MeanIoU, iNat &lt; PxAcc, Randall</b>	<b>-8.581</b>	< .001	<b>-1.108</b>
<b>MeanIoU, Randall &lt; PxAcc, iNat</b>	<b>-21.446</b>	< .001	<b>-2.769</b>
<b>MeanIoU, Randall &lt; PxAcc, Randall</b>	<b>-21.678</b>	< .001	<b>-2.799</b>
<b>PxAcc, iNat &gt; PxAcc, Randall</b>	<b>6.657</b>	< .001	<b>0.859</b>

*Note.* Student's  $t$ -test. Bonferroni-corrected (for multiple comparisons) post hoc paired-samples  $t$ -tests comparing the four image segmentation validation metrics; pixel accuracy (PxAcc, Eq. 1), mean accuracy (MeanAcc, Eq. 2), mean intersection over union (MeanIoU, Eq. 3), and frequency weighted intersection over union (FreqIoU, Eq. 4), from either iNaturalist (iNat) or J.E. Randall's fish photos collection (Randall). Cohen's  $d$  does not correct for multiple comparisons. Significant comparisons are bolded

## **Color Pattern Analysis Comparison**

From an ecological and evolutionary perspective, color pattern conspicuousness plays a critical role in the diversification of visual systems, signaling, and defensive behavior (Barlow 1972; Neudecker 1989; Domeier & Colin 1997; Marshall 2000; Losey *et al.* 2003; Marshall *et al.* 2003a; Marshall *et al.* 2003b; Randall 2005; Cheney *et al.* 2009; Salis *et al.* 2018), in addition to speciation processes (Bellwood *et al.* 2015; Bellwood, Goatley & Bellwood 2017; Salis *et al.* 2018; Alfaro *et al.* 2019; Hemingson *et al.* 2019; Salis *et al.* 2019). To better understand the evolution and divergence of organismal coloration and patterning, color pattern analysis has become increasingly common in ecological and evolutionary studies due to conceptual advances and the availability of new software tools (Endler 2012; Maia *et al.* 2013; Endler, Cole & Kranz 2018; Van Belleghem *et al.* 2018; Maia *et al.* 2019; Weller & Westneat 2019; Van Den Berg *et al.* 2020). These software approaches typically require input images with transparent (or artificially/uniformly colored) backgrounds to ensure that color pattern metrics are not influenced by similar, overlapping background pixels. As such, a common first step in color pattern analysis is image segmentation. This step is usually performed by hand, creating a potential bottleneck for larger scale analyses of color pattern.

Here, we compared an analysis of butterflyfish color pattern that used manual image segmentation from prior work (Alfaro *et al.* 2019) and compared color pattern geometry metrics of these manually segmented reef fishes to automatically segmented images of the same fishes with the custom fish segmentation model presented with *Sashimi*. Ninety-six of J.E. Randall's images of butterflyfishes were obtained from the Bishop Museum (<http://pbs.bishopmuseum.org/images/JER/>) and were segmented both manually by experts and automatically by *Sashimi*. Color pattern analyses were subsequently conducted in *pavo* (Version

2.0.0; Maia *et al.* 2013; Maia *et al.* 2019). For color pattern quantification, images were subsampled using a  $100 \times 100$  pixel grid and RGB values were used to compute color distances within each color region as a proxy for photoreceptor curves and spectral data. Two color measurements (Eqs. S1, S2) and luminance (Eq. S3) were computed for each color region according to Endler (2012) and were then used to compute Euclidean distances between regions to estimate chromatic and achromatic boundary strengths. The color pattern geometry variables computed and used in analyses here were: overall transition density ( $m$ ), aspect ratio ( $A$ ), scaled Simpson color class diversity ( $J_c$ ), scaled Simpson transition diversity ( $J_t$ ), and mean chromatic ( $m\_dS$ ) and achromatic ( $m\_dL$ ) boundary strength.

$$\frac{R - G}{R + G} \quad (S1)$$

$$\frac{G - B}{G + B} \quad (S2)$$

$$R + G + B \quad (S3)$$

## Results

We compared the results of color pattern geometry variables from an earlier study (Alfaro *et al.* 2019) using manual and *Sashimi* segmented images with a multivariate analysis of variance (MANOVA) to test whether the two image segmentation workflows generated similar values for the color pattern geometry descriptors. We found no significant effect of method of generating the segmented image (manually by hand, automatically by deep learning) for any of the color pattern geometry variables (all  $p > .05$ ; bolded in Table S4). Descriptive statistics for each color pattern

geometry variable for images segmented manually by hand and automatically by deep learning are presented in Table S5.

**Table S4****MANOVA: Pillai Test**

Cases	df	Approx. F	Trace Pillai	Num df	Den df	p
(Intercept)	1	7034.802	0.996	6	185.000	< .001
<b>method</b>	<b>1</b>	<b>1.708</b>	<b>0.052</b>	<b>6</b>	<b>185.000</b>	<b>0.121</b>
Residuals	190					

**Follow-up ANOVAs****ANOVA: *m***

Cases	Sum of Squares	df	Mean Square	F	p
(Intercept)	3.457	1	3.457	1810.285	< .001
<b>method</b>	<b>0.005</b>	<b>1</b>	<b>0.005</b>	<b>2.605</b>	<b>0.108</b>
Residuals	0.363	190	0.002		

**ANOVA: *A***

Cases	Sum of Squares	df	Mean Square	F	p
(Intercept)	245.472	1	245.472	4760.218	< .001
<b>method</b>	<b>0.010</b>	<b>1</b>	<b>0.010</b>	<b>0.203</b>	<b>0.653</b>
Residuals	9.798	190	0.052		

**ANOVA: *Jc***

Cases	Sum of Squares	df	Mean Square	F	p
(Intercept)	75.000	1	75.000	9075.140	< .001
<b>method</b>	<b>0.027</b>	<b>1</b>	<b>0.027</b>	<b>3.320</b>	<b>0.070</b>
Residuals	1.570	190	0.008		

**ANOVA: *Jt***

Cases	Sum of Squares	df	Mean Square	F	p
(Intercept)	70.664	1	70.664	6747.701	< .001
<b>method</b>	<b>0.011</b>	<b>1</b>	<b>0.011</b>	<b>1.078</b>	<b>0.300</b>
Residuals	1.990	190	0.010		

**ANOVA: *m\_dS***

Cases	Sum of Squares	df	Mean Square	F	p
(Intercept)	3.086	1	3.086	640.887	< .001
<b>method</b>	<b>7.083e -5</b>	<b>1</b>	<b>7.083e -5</b>	<b>0.015</b>	<b>0.904</b>
Residuals	0.915	190	0.005		

**ANOVA: *m\_dL***

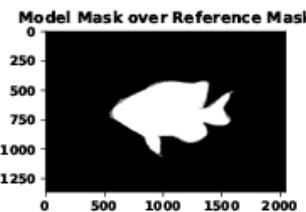
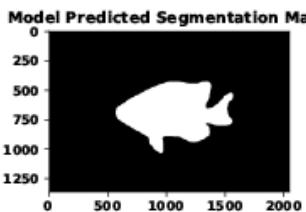
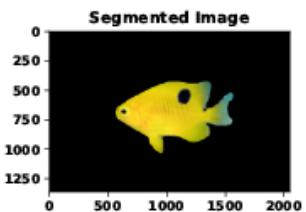
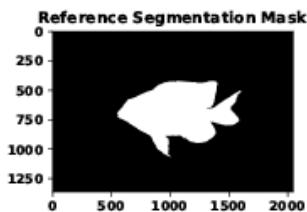
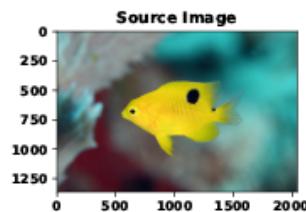
Cases	Sum of Squares	df	Mean Square	F	p
(Intercept)	166.760	1	166.760	7280.861	< .001
<b>method</b>	<b>0.015</b>	<b>1</b>	<b>0.015</b>	<b>0.641</b>	<b>0.424</b>
Residuals	4.352	190	0.023		

**Table S5****Descriptive Statistics for Color Pattern Geometry Variables**

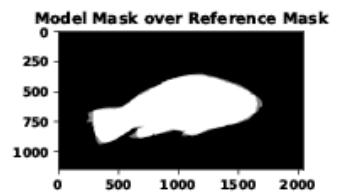
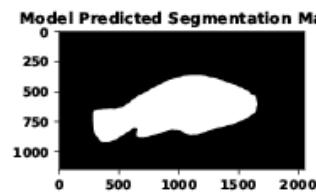
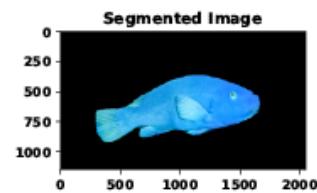
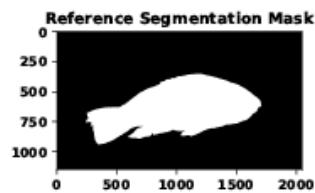
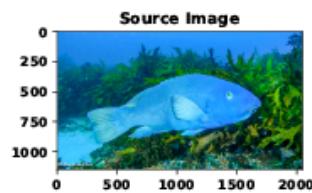
	m		A		Jc		Jt		m_dS		m_dL	
	deep_learning	human										
Valid	96	96	96	96	96	96	96	96	96	96	96	96
Missing	0	0	0	0	0	0	0	0	0	0	0	0
Mean	0.129	0.139	1.123	1.138	0.613	0.637	0.599	0.614	0.127	0.126	0.923	0.941
Std. Deviation	0.043	0.045	0.233	0.221	0.089	0.093	0.102	0.102	0.070	0.069	0.149	0.154
Minimum	0.037	0.053	0.633	0.685	0.392	0.411	0.349	0.366	0.016	0.018	0.606	0.594
Maximum	0.290	0.306	1.774	1.776	0.758	0.801	0.854	0.857	0.411	0.368	1.215	1.268

## **iNaturalist Novel Test Set ( $n = 30$ ) Visual Evaluations**

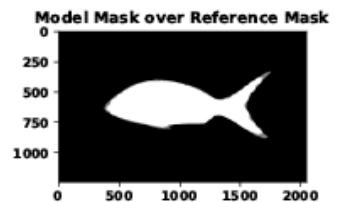
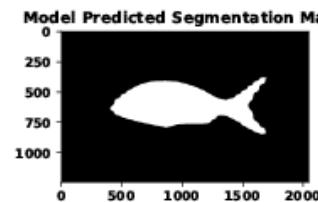
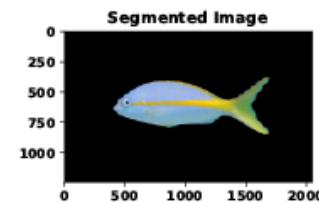
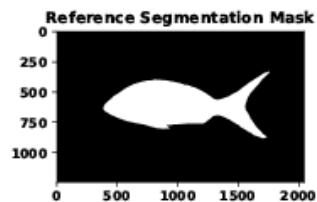
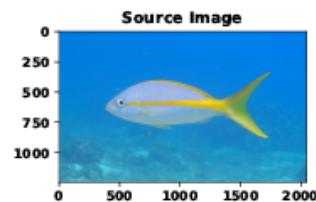
147548567



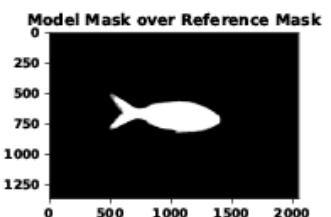
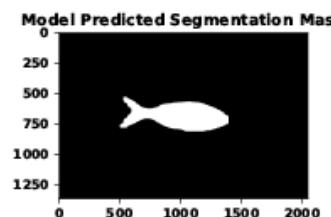
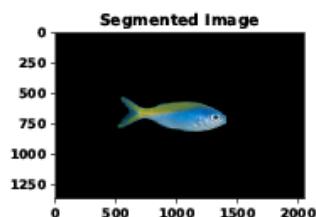
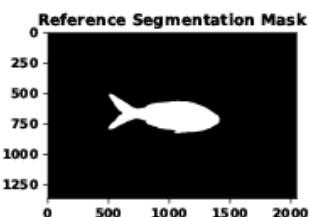
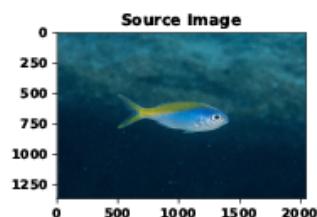
148679170



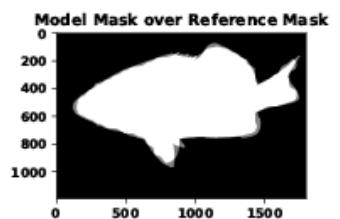
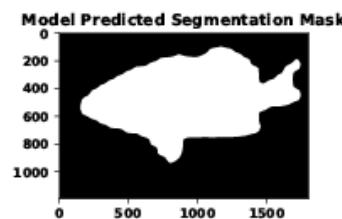
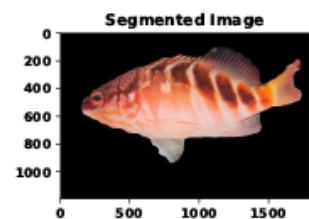
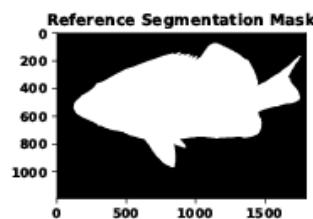
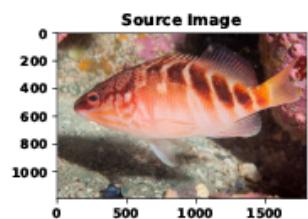
149293658



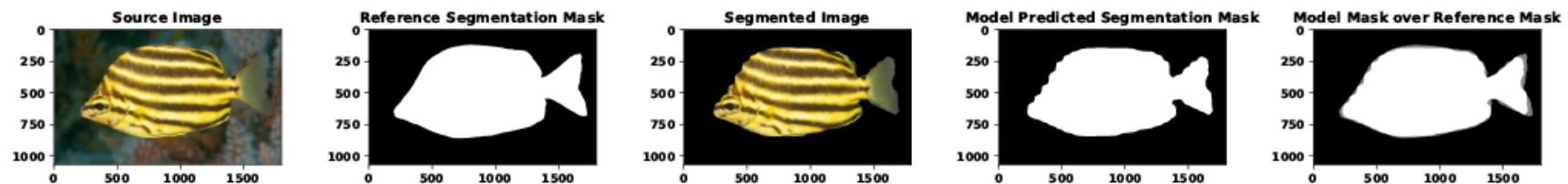
149359148



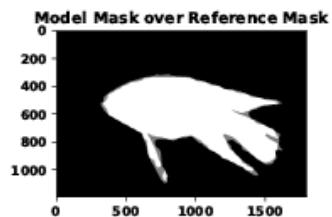
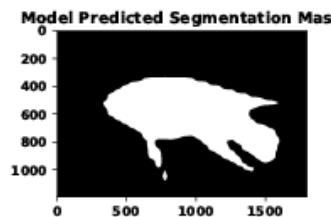
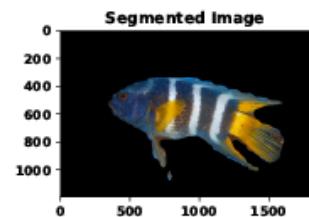
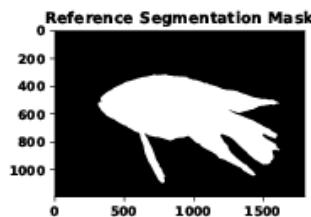
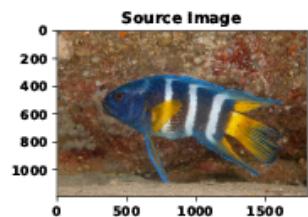
149819354



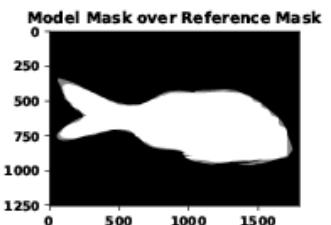
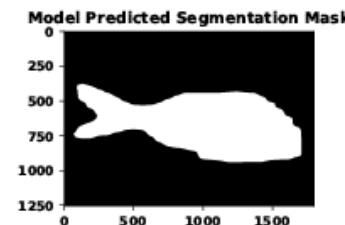
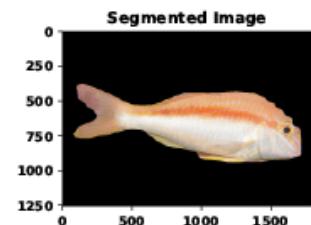
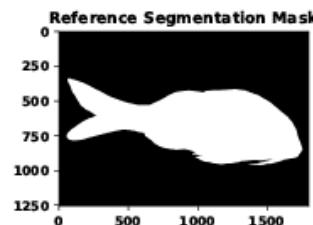
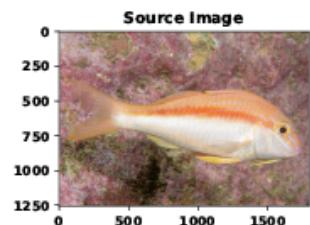
149819548



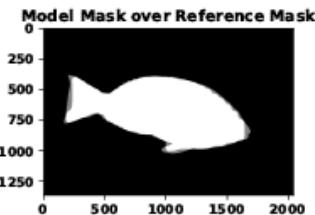
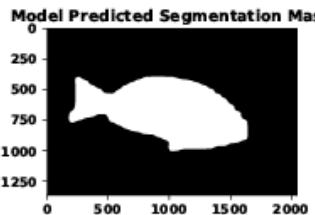
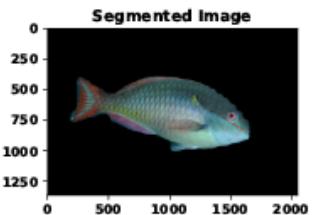
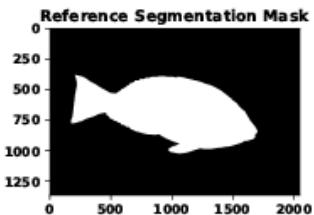
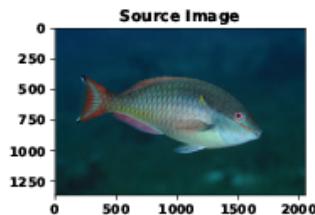
149819653



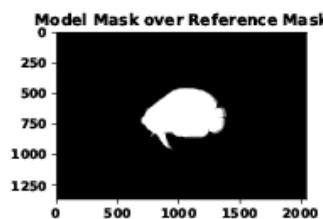
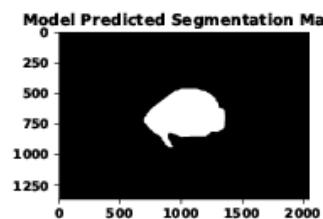
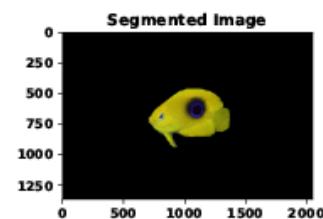
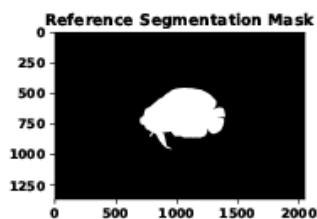
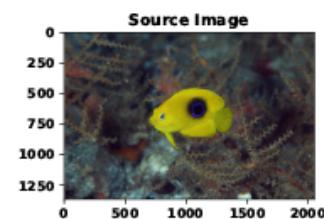
149819703



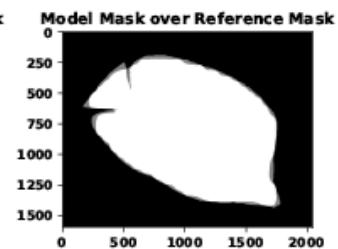
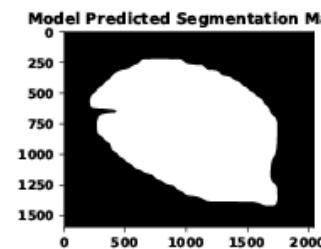
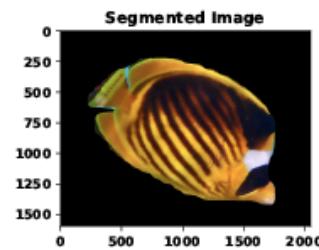
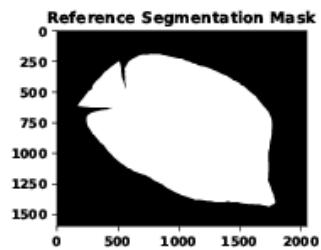
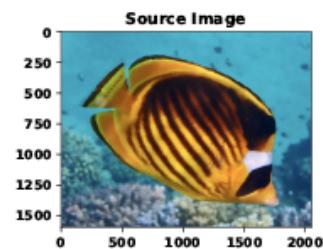
149996859



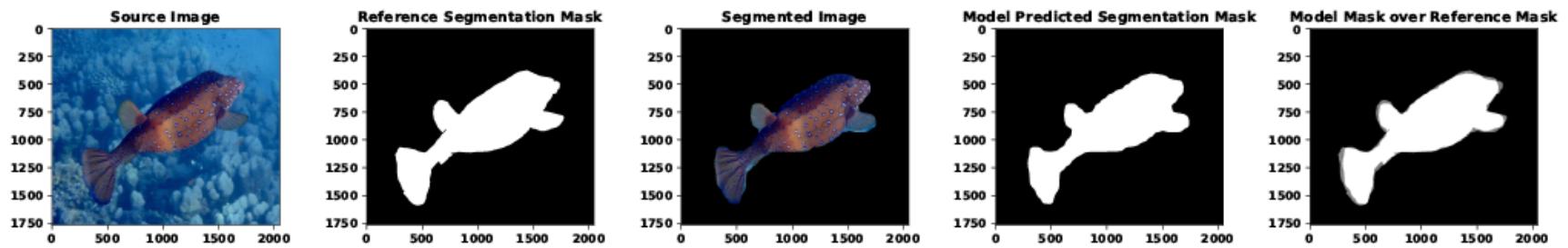
150197026



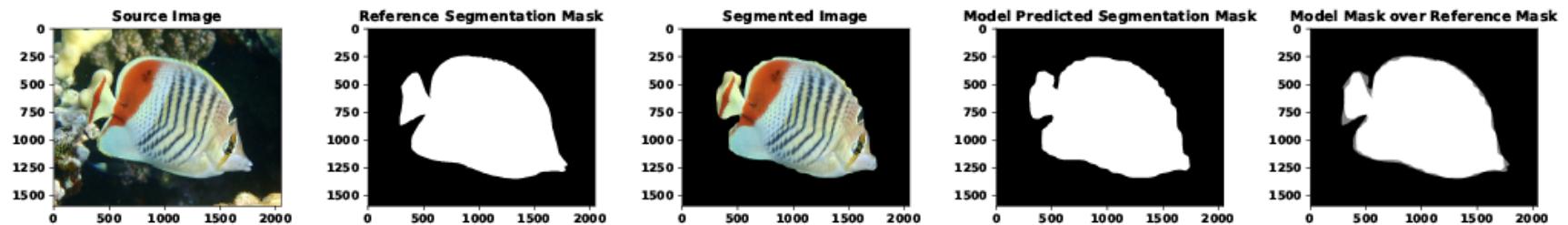
150210847



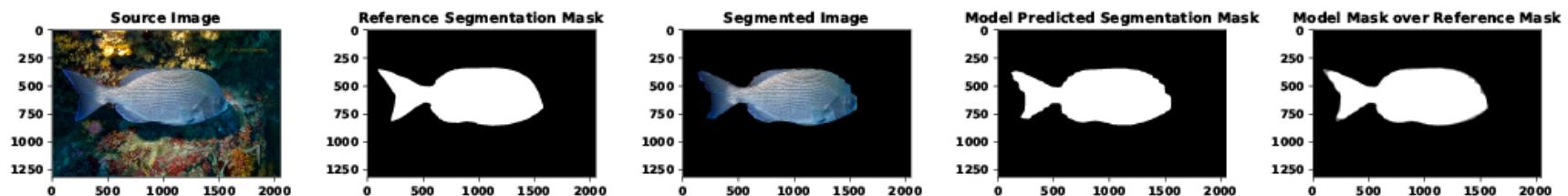
150241404



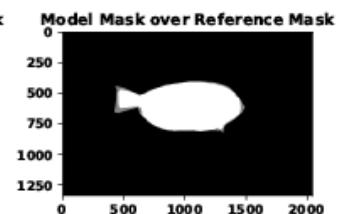
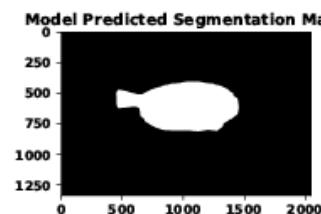
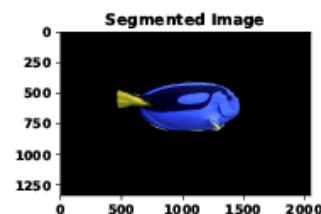
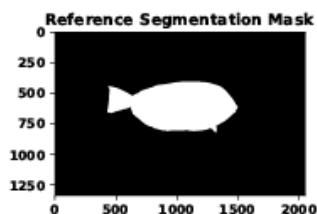
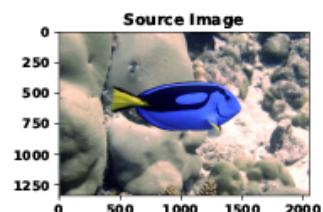
150452375



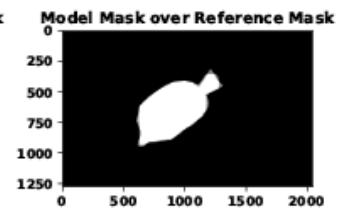
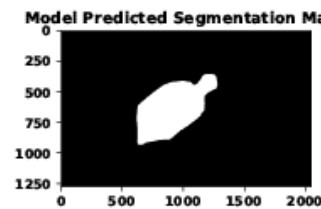
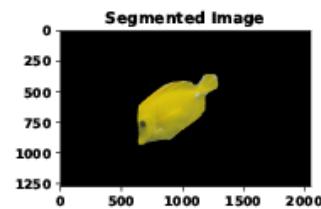
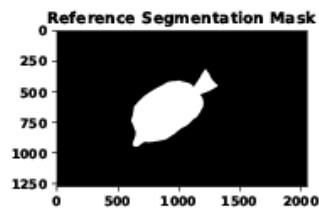
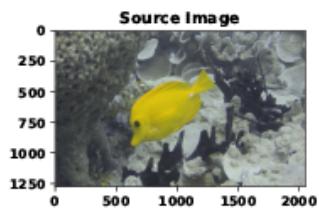
150508414



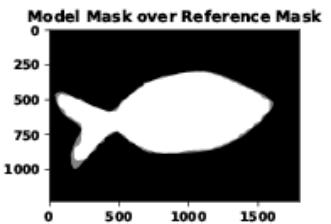
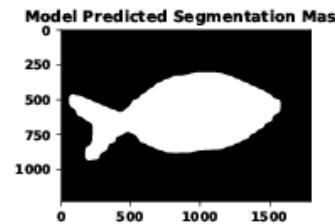
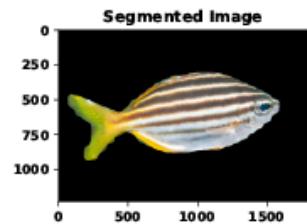
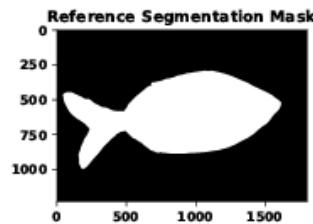
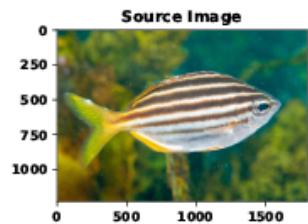
150618936



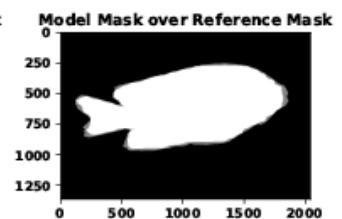
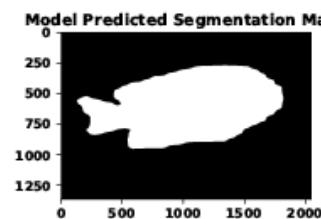
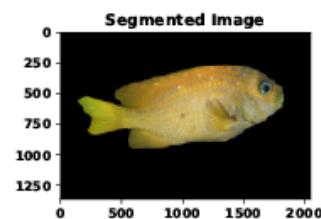
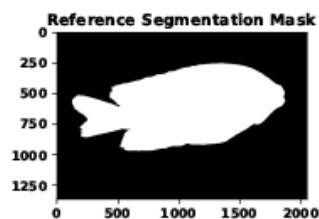
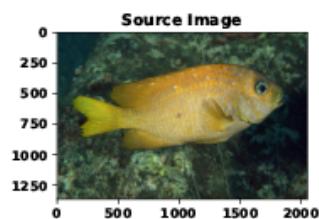
150620142



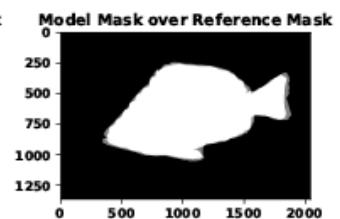
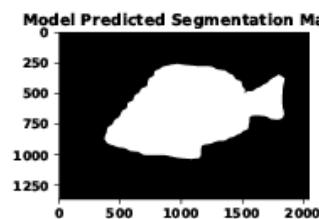
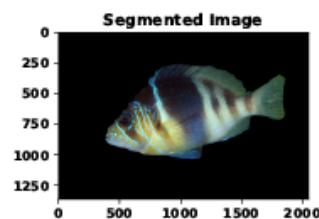
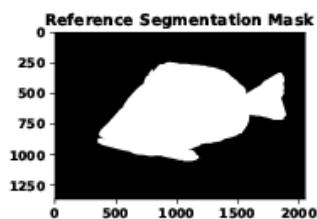
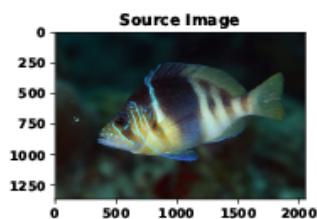
150623552



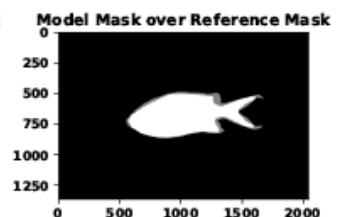
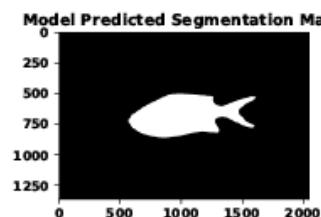
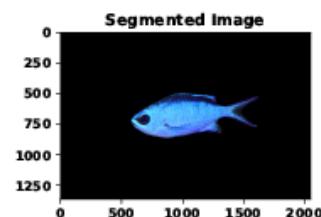
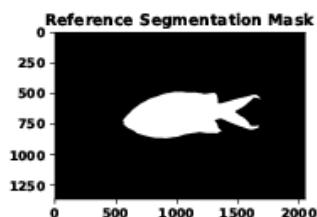
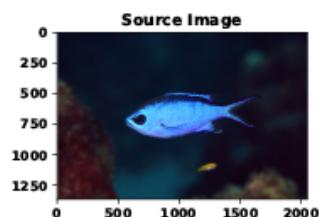
150719728



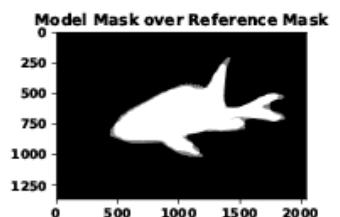
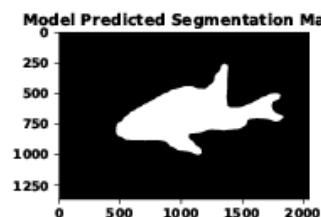
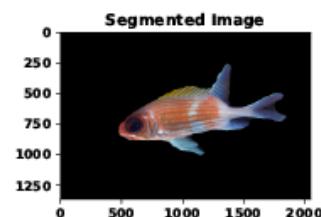
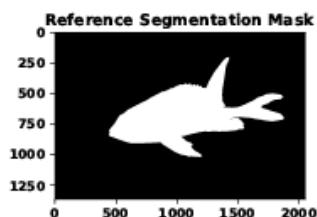
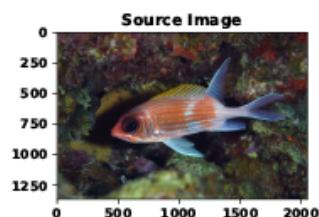
150720200



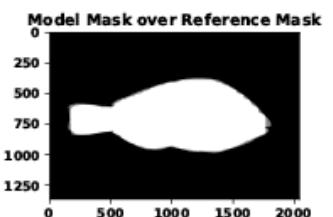
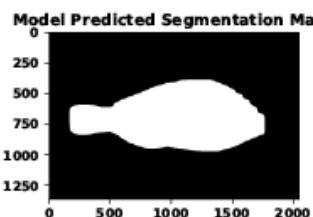
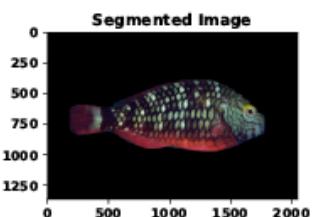
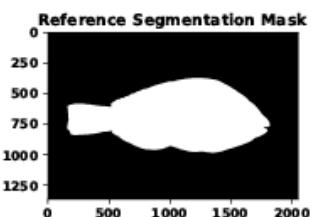
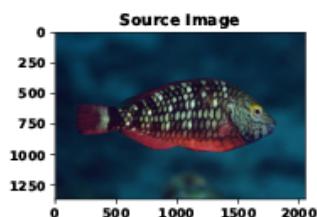
150720956



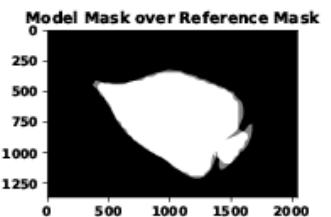
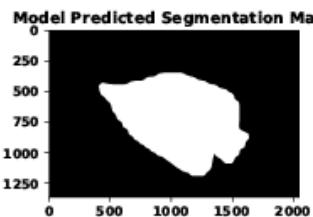
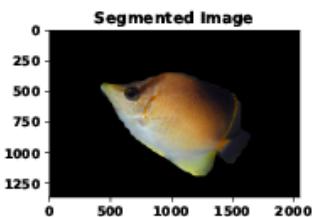
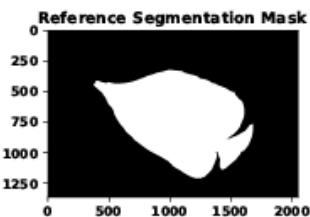
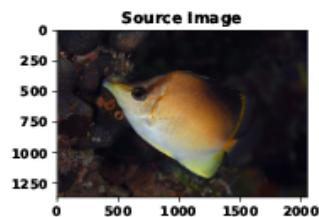
150721261



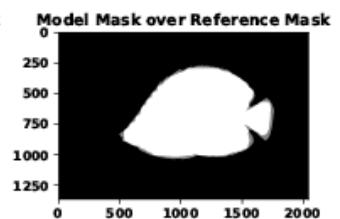
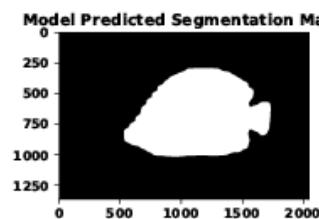
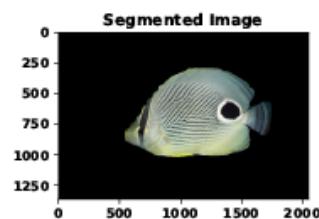
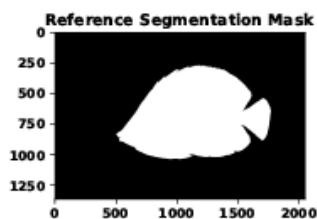
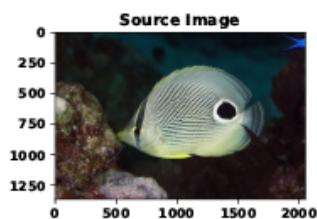
150721405



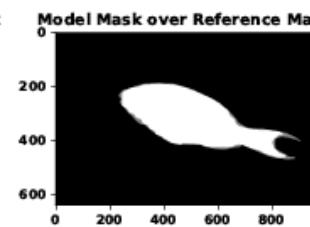
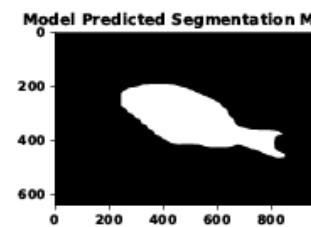
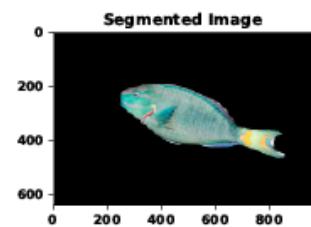
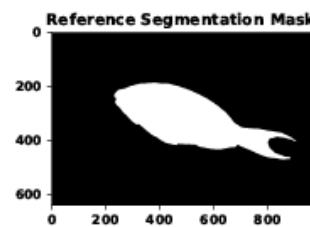
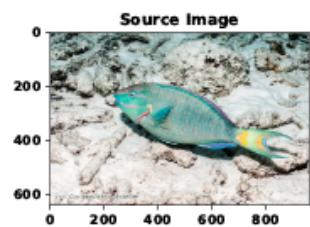
150724897



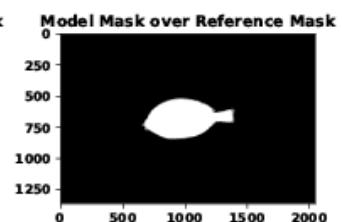
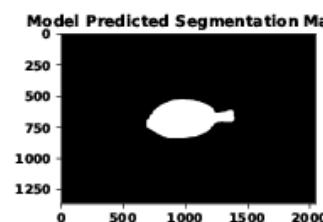
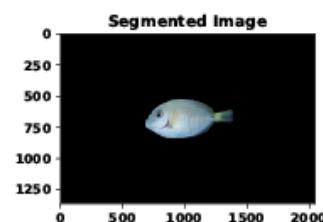
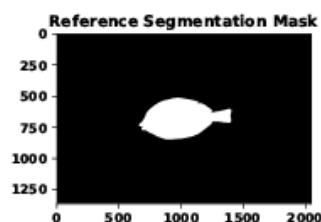
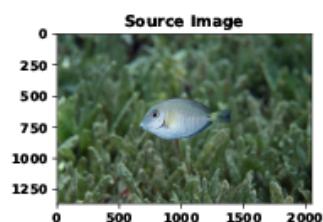
150725386



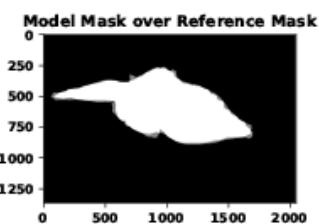
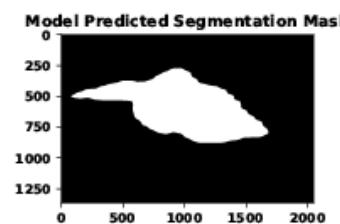
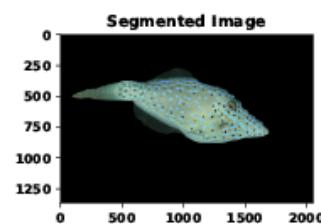
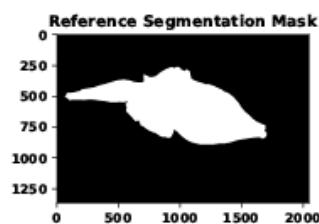
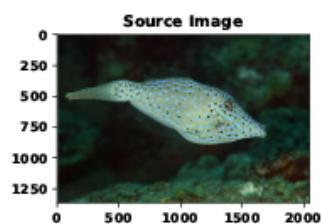
150741522



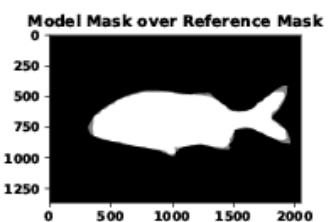
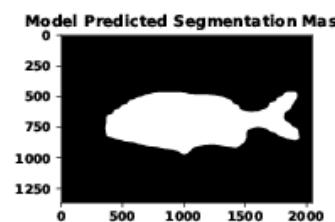
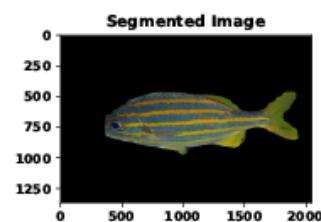
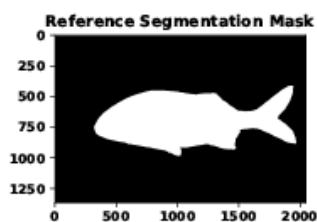
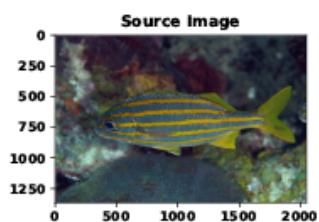
150841882



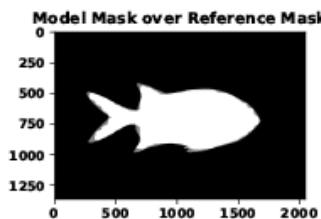
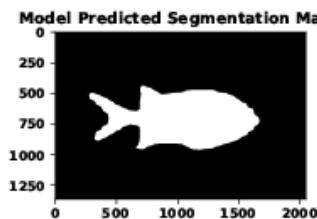
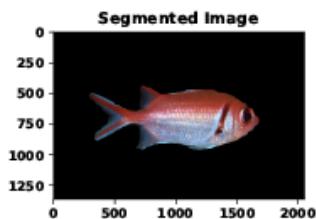
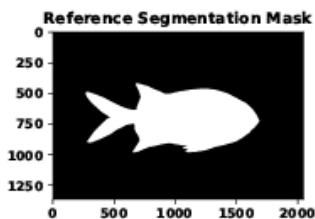
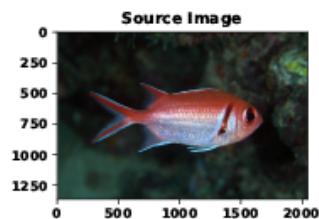
150843014



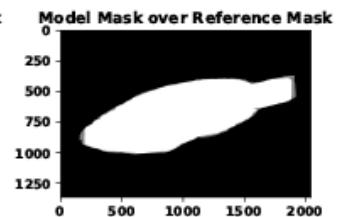
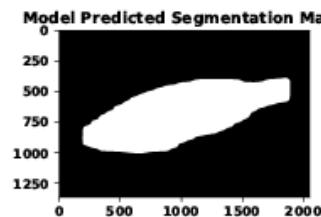
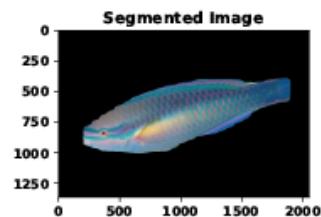
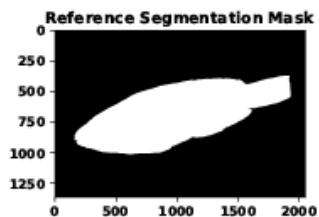
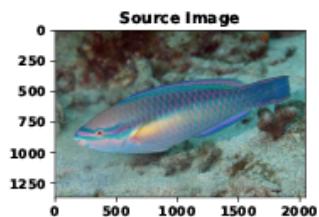
150843427



150843567

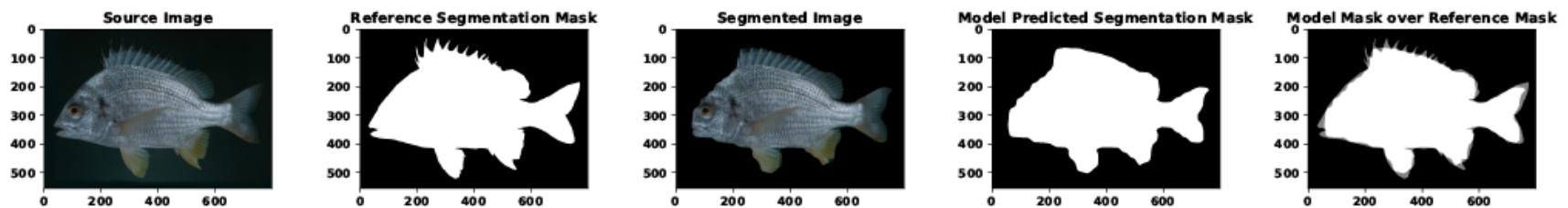


150845262

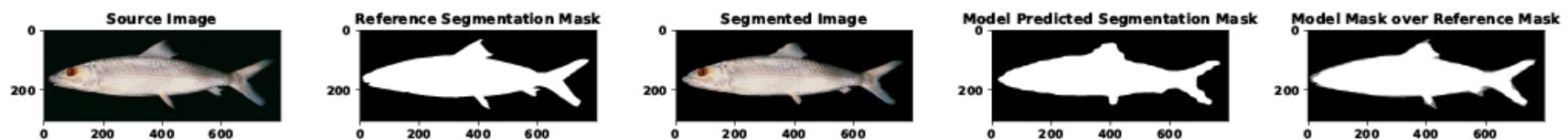


**J.E. Randall Novel Test Set ( $n = 30$ ) Visual Evaluations**

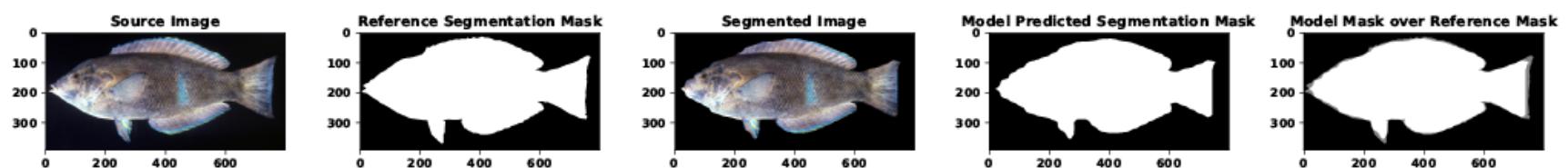
*Acanthopagrus* *Acanthopagrus latus* 705783125



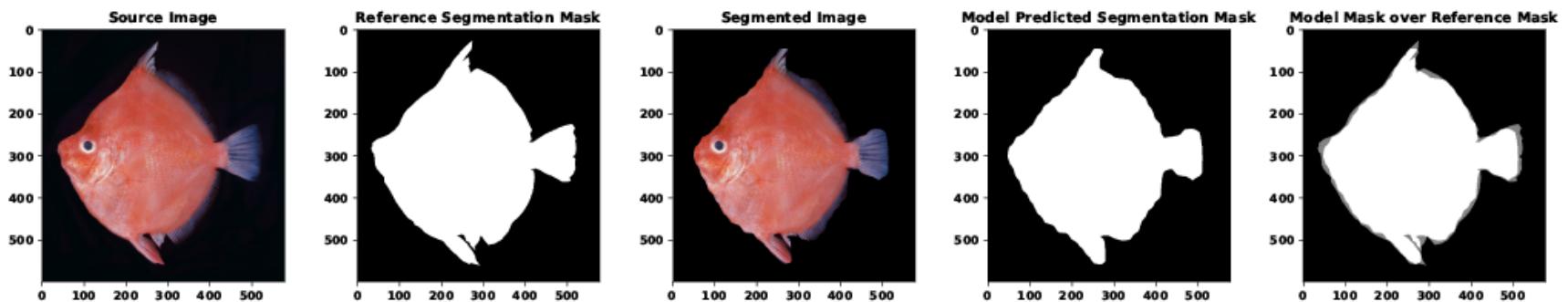
Albulidae\_Albulidae\_glossodonta\_575984577



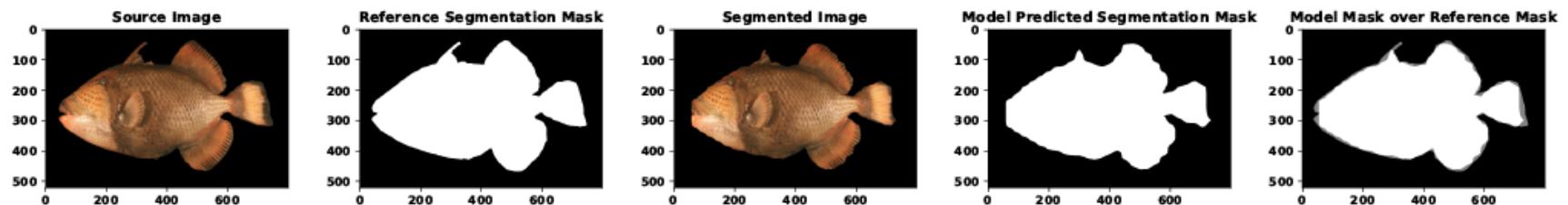
Anampsese\_Anampsese geographicus\_1054869801



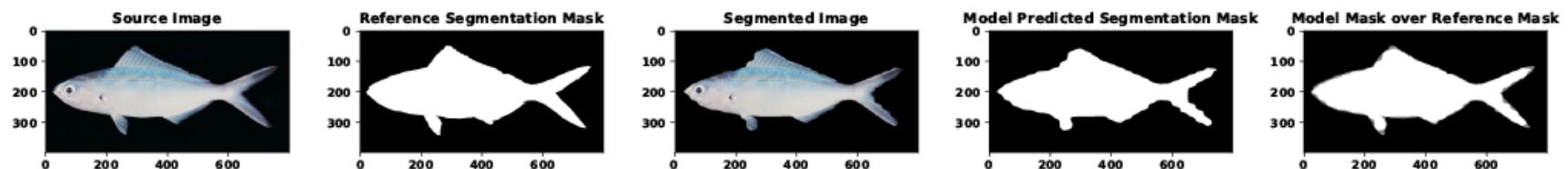
Antigonia\_Antigonia capros\_-1671622598



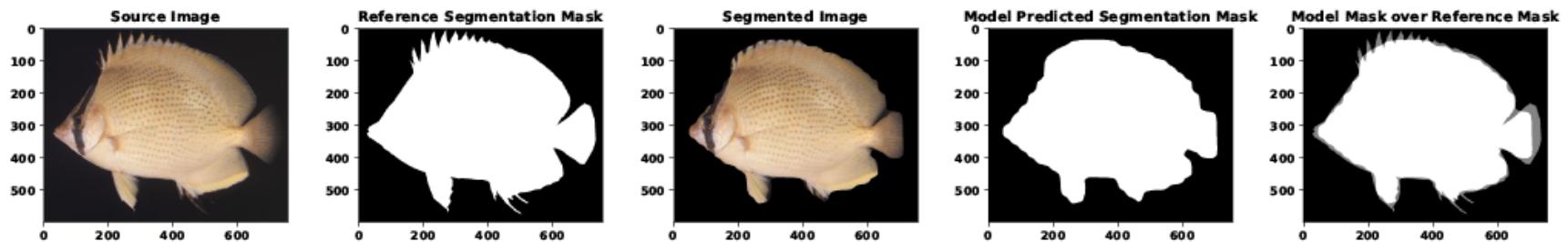
*Ballistoides* *Ballistoides viridescens* 1486707515



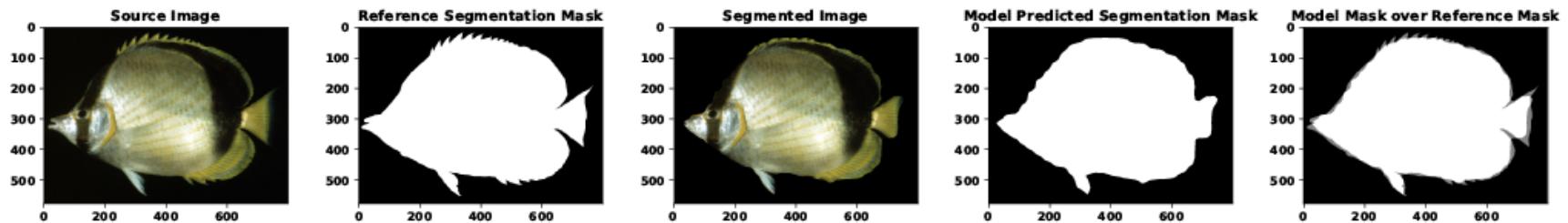
*Caesio* \_*Caesio caeruleaurea* \_997393396



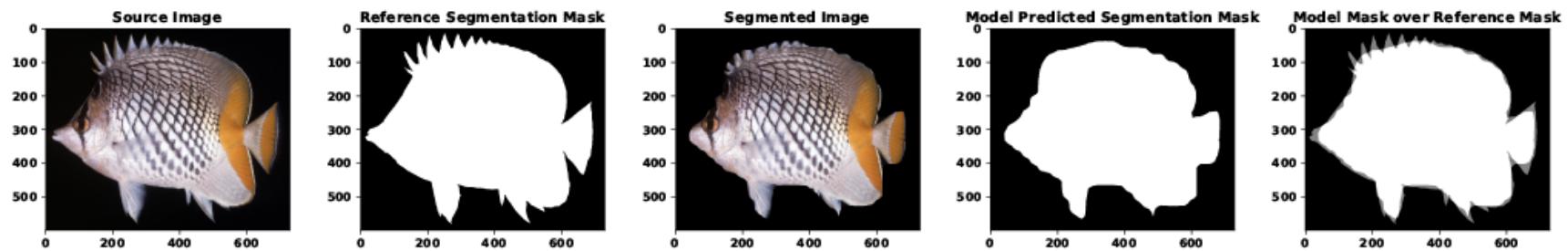
Chaetodon\_Chaetodon citrinellus\_-1736280642



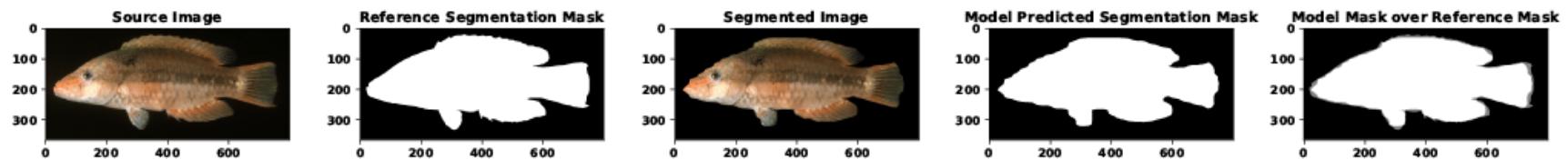
*Chaetodon* \_*Chaetodon gardineri*\_ -1706450280



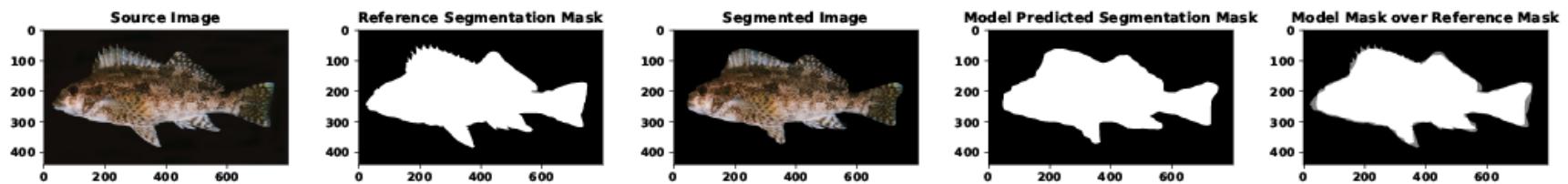
*Chaetodon* \_ *Chaetodon xanthurus* \_858508951



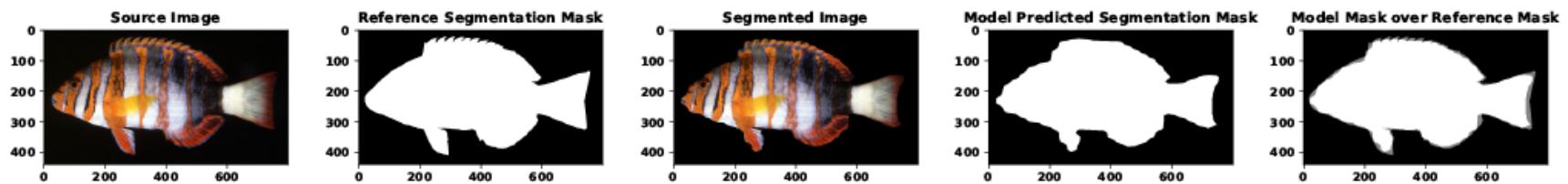
Cheilinus Cheilinus mentalis\_1645979501



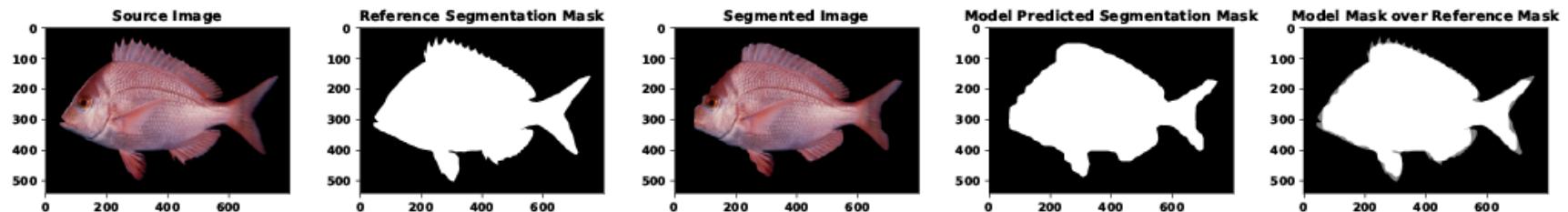
Chironemus\_ Chironemus marmoratus \_-1819846277



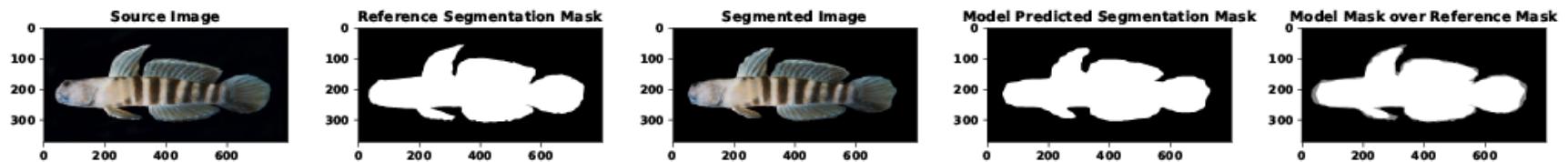
*Choerodon Choerodon fasciatus* -618230879



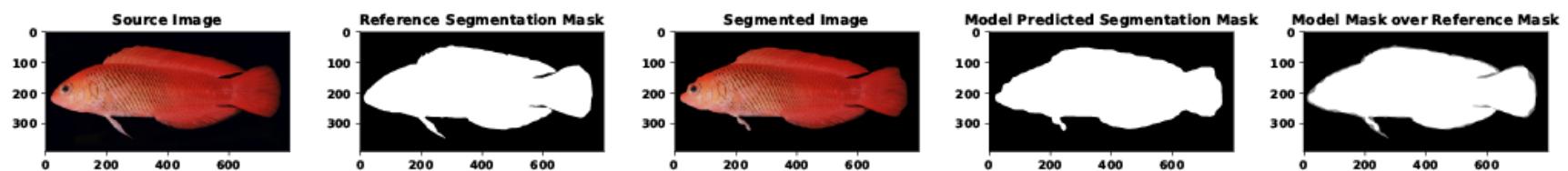
*Chrysoblephus* *Chrysoblephus puniceus* 2008107775



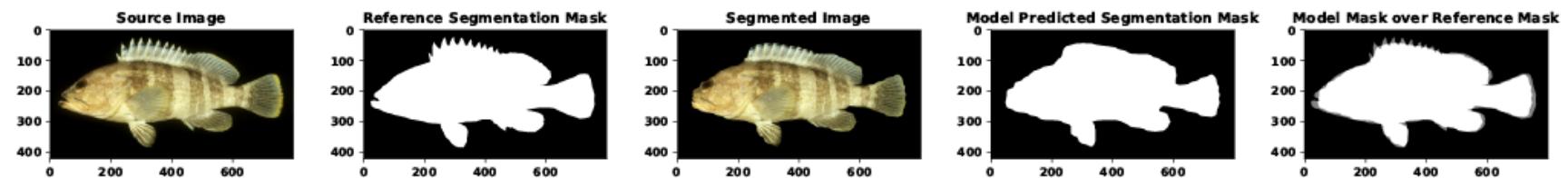
Cryptocentrus\_Cryptocentrus lutheri\_662979441



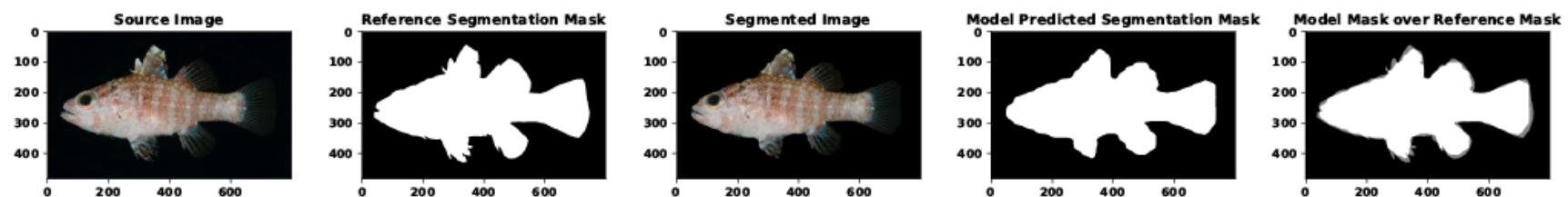
Cypho\_Cypho purpurascens\_678471231



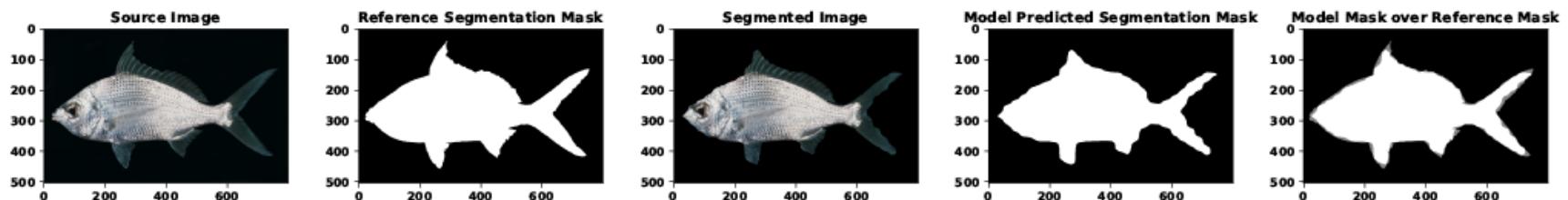
*Epinephelus* *Epinephelus awoara* \_352109216



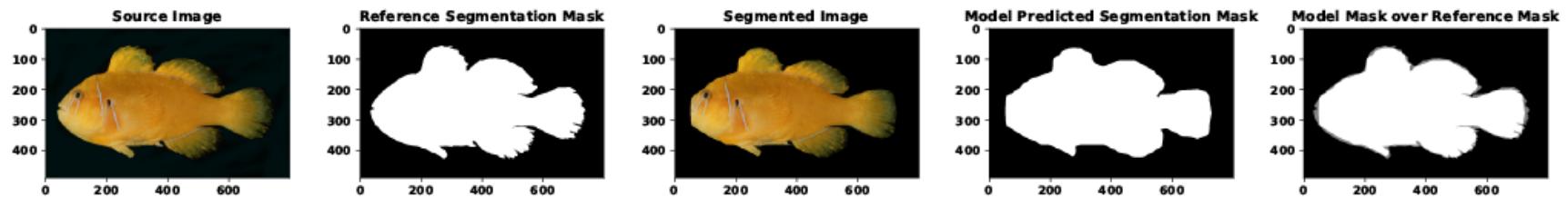
*Fowleria* *Fowleria valulae* \_1990154993



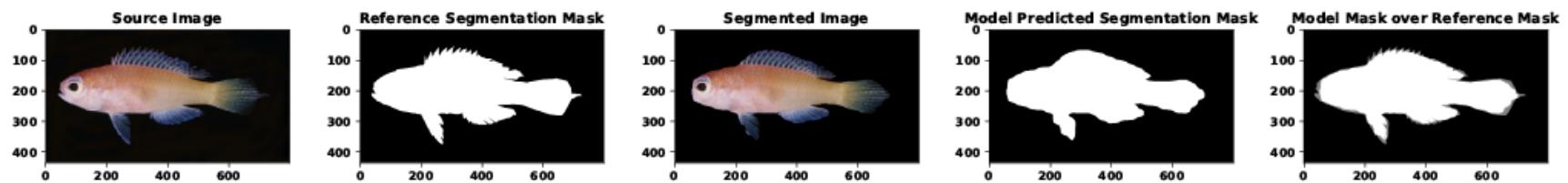
*Gerres Gerres longirostris -266959284*



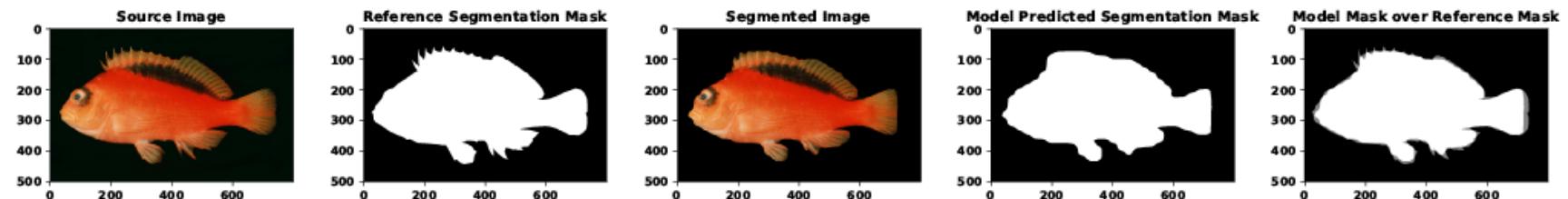
Gobiodon\_Gobiodon citrinus\_1046390725



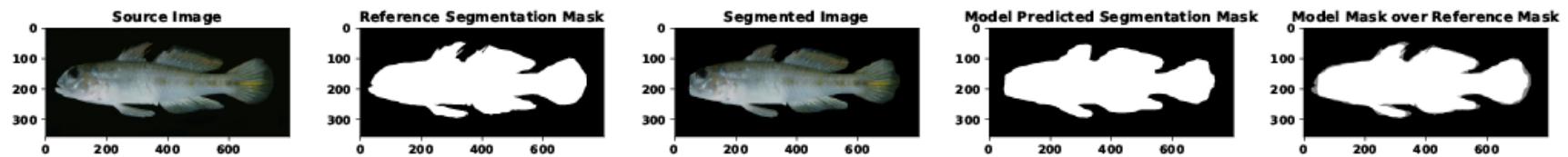
Grammatonotus\_Grammatonotus laysanus\_169863020



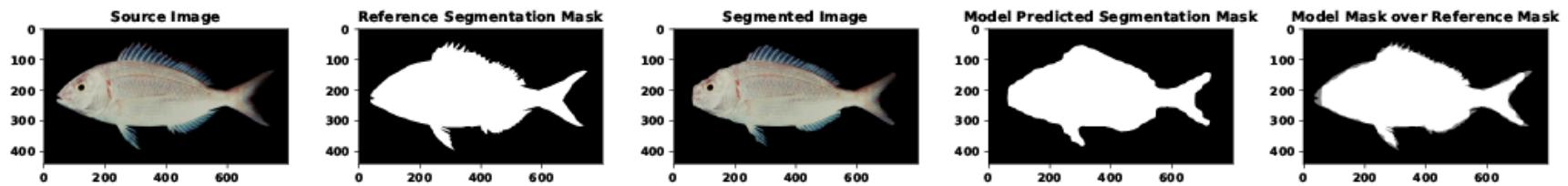
*Neocirrhitest* *Neocirrhitest armatus* \_1606479153



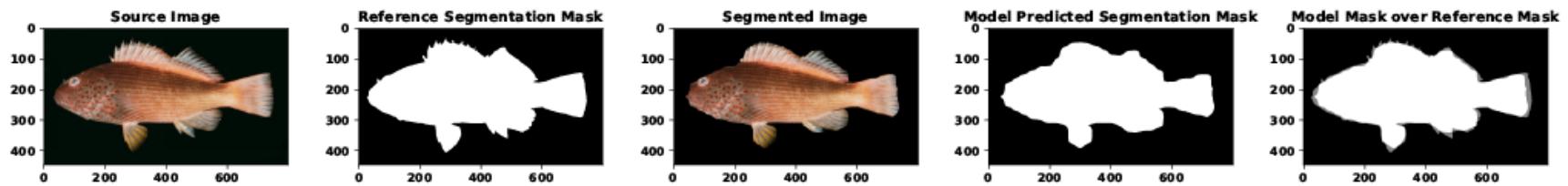
Olopomus\_Olopomus olopomus\_987273204



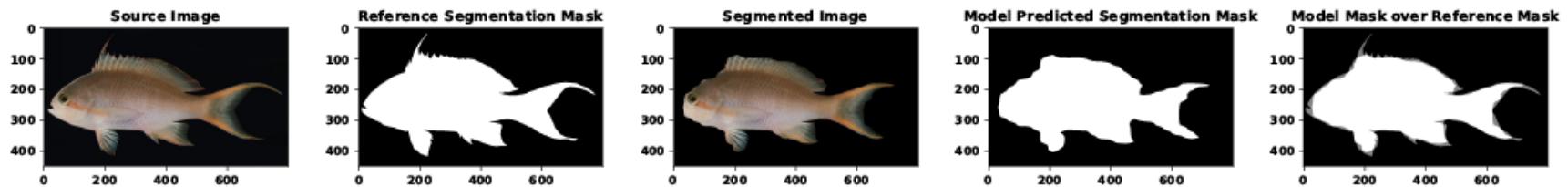
*Pagellus* *Pagellus affinis* 511019768



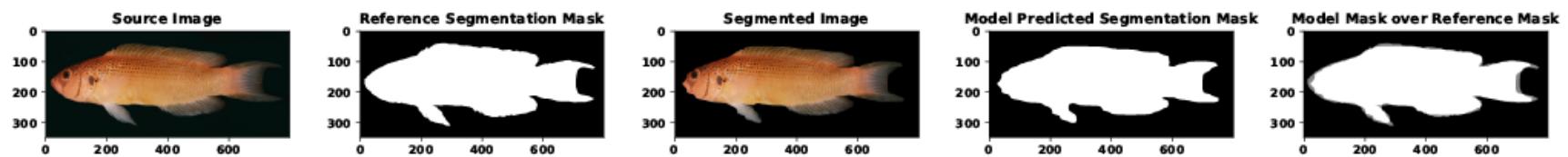
*Paracirrhites* *Paracirrhites forsteri* -270789107



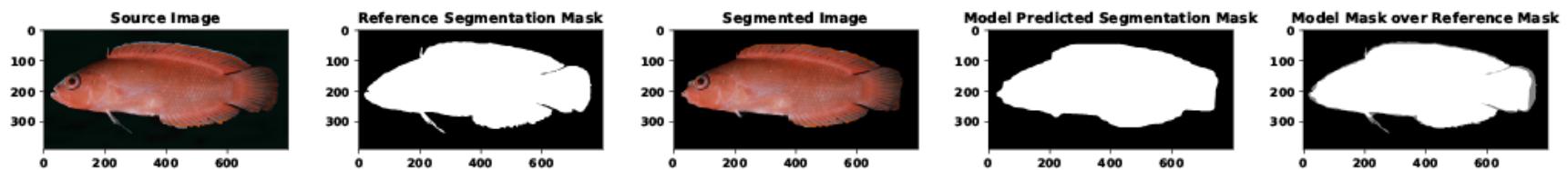
Pseudanthias\_Pseudanthias huchti\_896234028



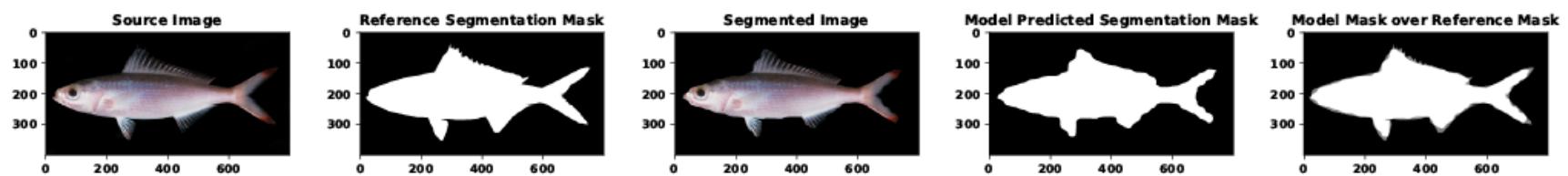
Pseudochromis\_Pseudochromis moorei\_-2131341876



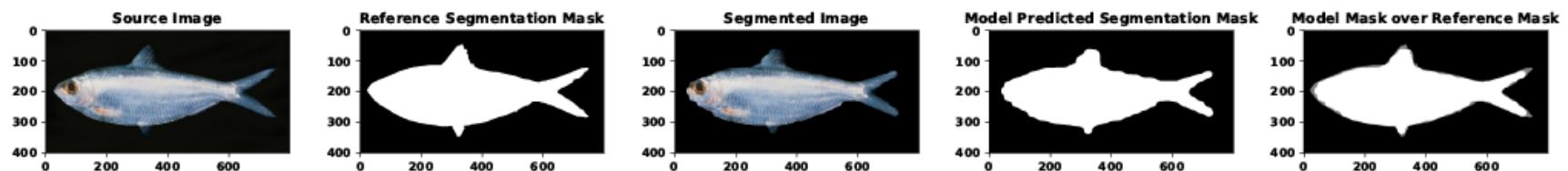
Pseudoplesiops\_Pseudoplesiops typus\_-1769782320



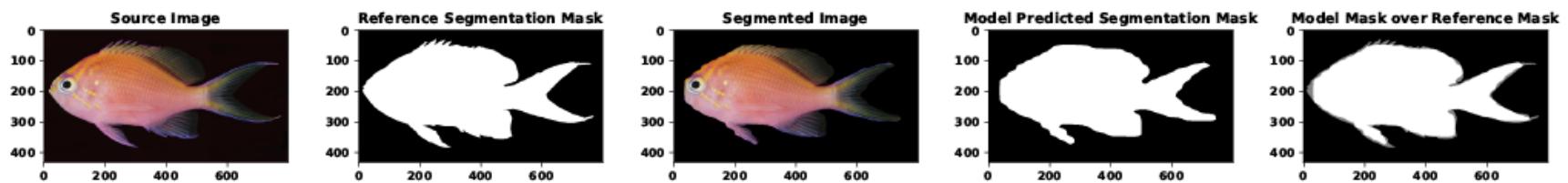
Pteroacesio\_Pteroacesio pisang\_-1779285115



Sardinella\_Sardinella albellus\_445765935



*Serranocirrhitus* *Serranocirrhitus latus* \_1070832338



## References

- Alfaro, M.E., Karan, E.A., Schwartz, S.T. & Shultz, A.J. (2019) The Evolution of Color Pattern in Butterflyfishes (Chaetodontidae). *Integr Comp Biol*, **59**, 604-615.
- Barlow, G.W. (1972) The attitude of fish eye-lines in relation to body shape and to stripes and bars. *Copeia*, 4-12.
- Bellwood, D.R., Goatley, C.H., Cowman, P.F. & Bellwood, O. (2015) The evolution of fishes on coral reefs: fossils, phylogenies and functions. *Ecology of Fishes on Coral Reefs*, 55-63.
- Bellwood, D.R., Goatley, C.H.R. & Bellwood, O. (2017) The evolution of fishes and corals on reefs: form, function and interdependence. *Biological Reviews*, **92**, 878-901.
- Cheney, K.L., Grutter, A.S., Blomberg, S.P. & Marshall, N.J. (2009) Blue and Yellow Signal Cleaning Behavior in Coral Reef Fishes. *Current Biology*, **19**, 1283-1287.
- Domeier, M.L. & Colin, P.L. (1997) Tropical reef fish spawning aggregations: defined and reviewed. *Bulletin of Marine Science*, **60**, 698-726.
- Endler, J.A. (2012) A framework for analysing colour pattern geometry: adjacent colours. *Biological Journal of the Linnean Society*, **107**, 233-253.
- Endler, J.A., Cole, G.L. & Kranz, A.M. (2018) Boundary strength analysis: Combining colour pattern geometry and coloured patch visual properties for use in predicting behaviour and fitness. *Methods in Ecology and Evolution*, **9**, 2334-2348.
- Hemingson, C.R., Cowman, P.F., Hodge, J.R. & Bellwood, D.R. (2019) Colour pattern divergence in reef fish species is rapid and driven by both range overlap and symmetry. *Ecology letters*, **22**, 190-199.

Losey, G., McFarland, W., Loew, E., Zamzow, J., Nelson, P. & Marshall, N. (2003) Visual biology of Hawaiian coral reef fishes. I. Ocular transmission and visual pigments. *Copeia*, **2003**, 433-454.

Maia, R., Eliason, C.M., Bitton, P.-P., Doucet, S.M. & Shawkey, M.D. (2013) pavo: an R package for the analysis, visualization and organization of spectral data. *Methods in Ecology and Evolution*, n/a-n/a.

Maia, R., Gruson, H., Endler, J.A. & White, T.E. (2019) pavo 2: new tools for the spectral and spatial analysis of colour in R. *Methods in Ecology and Evolution*, **10**, 1097-1107.

Marshall, N., Jennings, K., McFarland, W., Loew, E. & Losey, G. (2003a) Visual biology of Hawaiian coral reef fishes. II. Colors of Hawaiian coral reef fish. *Copeia*, **2003**, 455-466.

Marshall, N., Jennings, K., McFarland, W., Loew, E. & Losey, G. (2003b) Visual biology of Hawaiian coral reef fishes. III. Environmental light and an integrated approach to the ecology of reef fish vision. *Copeia*, **2003**, 467-480.

Marshall, N.J. (2000) Communication and camouflage with the same 'bright' colours in reef fishes. *Philos Trans R Soc Lond B Biol Sci*, **355**, 1243-1248.

Neudecker, S. (1989) Eye camouflage and false eyespots: chaetodontid responses to predators. *The butterflyfishes: success on the coral reef*, pp. 143-158. Springer.

Randall, J.E. (2005) A review of mimicry in marine fishes. *ZOOLOGICAL STUDIES-TAIPEI-*, **44**, 299.

Salis, P., Lorin, T., Laudet, V. & Frédéric, B. (2019) Magic Traits in Magic Fish: Understanding Color Pattern Evolution Using Reef Fish. *Trends in Genetics*, **35**, 265-278.

Salis, P., Roux, N., Soulat, O., Lecchini, D., Laudet, V. & Frédéric, B. (2018) Ontogenetic and phylogenetic simplification during white stripe evolution in clownfishes. *BMC Biology*, **16**.

Van Belleghem, S.M., Papa, R., Ortiz-Zuazaga, H., Hendrickx, F., Jiggins, C.D., Owen Mcmillan, W. & Counterman, B.A. (2018) patternize: An R package for quantifying colour pattern variation. *Methods in Ecology and Evolution*, **9**, 390-398.

Van Den Berg, C.P., Troscianko, J., Endler, J.A., Marshall, N.J. & Cheney, K.L. (2020) Quantitative Colour Pattern Analysis (QCPA): A comprehensive framework for the analysis of colour patterns in nature. *Methods in Ecology and Evolution*, **11**, 316-332.

Weller, H.I. & Westneat, M.W. (2019) Quantitative color profiling of digital images with earth mover's distance using the R package colordistance. *PeerJ*, **7**, e6398.