SUPPLEMENTARY TABLES

Table S1. Raw fish thumbnails training dataset

5

Classes	Number of thumbnails
Abudefduf sparoides	1241
Abudefduf vaigiensis	5674
Chaetodon trifascialis	1456
Chromis weberi	3576
Dascyllus carneus	2276
Lutjanus kasmira	1652
Monotaxis grandoculis	1239
Myripristis botche	1264
Naso elegans	2068
Mulloidichtys vanicolensis	1264
Naso vlamingii	1789
Nemateleotris magnifica	1189
Odonus niger	2986
Plectroglyphidodon lacrymatus	652
Pomacentrus sulfureus	5176
Preocaesio tile	3088
Pygoplytes diacanthus	1106
Thalassoma hardwicke	1579

Zanclus cornutus	1886
Zebrasoma scopas	1835

Table S2. The four thumbnails datasets used to train the four models, with for each the number of thumbnails per class in the training datasets, with class “environment” gathering thumbnails of water and substrate (sand, corals) while “Part of fish” gathers all thumbnails of half of a fish individual and the "part of species" classes contain thumbnails of half of individuals for each species.

Species	Only whole fish (T1)	Whole fish + “Part of fish” (T2)	Whole fish + Environment + “Part of fish” (T3)	Whole fish + environment + “Part of species” (T4)
Abudefduf sparoides	2482	2482	2482	2482
Abudefduf vaigiensis	11328	11328	11328	11328
Chaetodon trifascialis	2912	2912	2912	2912
Chromis weberi	7152	7152	7152	7152
Dascyllus carneus	4552	4552	4552	4552
Lutjanus kasmira	3300	3300	3300	3300
Monotaxis grandoculis	2478	2478	2478	2478
Mulloidichtys vanicolensis	2528	2528	2528	2528
Myripristis botche	2528	2528	2528	2528
Naso elegans	4138	4138	4138	4138
Naso vlamingii	3578	3578	3578	3578
Nemateleotris magnifica	2378	2378	2378	2378
Odonus niger	5972	5972	5972	5972
Plectroglyphidodon lacrymatus	1304	1304	1304	1304
Pomacentrus sulfureus	10352	10352	10352	10352
Preocaesio tile	6176	6176	6176	6176
Pygoplytes diacanthus	2212	2212	2212	2212
Thalassoma hardwicke	3158	3158	3158	3158
Zanclus cornutus	3772	3772	3772	3772
Zebrasoma scopas	3670	3670	3670	3670
Part of *Abudefduf sparoides*				4964
Part of *Abudefduf vaigiensis*				22656
Part of *Chaetodon trifascialis*				5824

Part of			14304
Naso elegans			
Part of			20704
Pomacentrus sulfureus			
Part of			6600
Lutjanus kasmira			
Part of			4424
Pygoplites diacanthus			
Part of			6316
Thalassoma hardwicke			
Part of			7544
Zanclus cornutus			
Part of			7340
Zebrasoma scopas			
Part of			14034
Chromis weberi			
Part of			4956
Monotaxis grandoculis			
Part of			1304
Plectroglyphidodon lacrymatus			
Part of			9097
Dascyllus carneus			
Part of			5056
Myripristis botche			
Part of			7156
Naso vlamingii			
Part of			4744
Nemateleotris magnifica			
Part of			11944
Odonus niger			
Part of			12352
Pterocaesio tile			
Part of			7528
Mulloidichtys vanicolensis			
Part of Fish	521555	521555	

Environment	862174	862174

Table S3: Number of thumbnails of each fish species present in test datasets used in this study

Class	Dataset for testing models performance	Dataset for testing model performance vs human performance
Abudefduf sparoides	103	88
Abudefduf vaigiensis	59	47
Chaetodon trifascialis	208	146
Chromis weberi	269	
Dascyllus carneus	269	
Monotaxis grandoculis	72	
Myripristis botche	20	
Naso elegans	189	165
Naso vlamingii	358	
Nemateleotris magnifica	246	
Odonus niger	176	
Plectroglyphidodon lacrymatus	150	
Pomacentrus sulfureus	1567	443
Pterocaesio tile	215	
Pygoplytes diacanthus	39	35
Thalassoma hardwicke	111	73
Zanclus cornutus	64	53
Zebrasoma scopas	184	144

	Total	4405	1197

Table S4. Performance of CNN model trained with T4 thumbnails set to identify nine fish species with no post processing; species are identified in columns and rows refer to whole fish and parts of fish present in the training dataset.

Part of species X means that some individual were recognized as part of a fish species.

Only percentages of over 1% are shown.

Species	*A.sparoides*	*A.vaigiensis*	*C.trifascialis*	*N.elegans*	*P.sulfureus*	*P.diacanthus*	*T.hardwicke*	*Z.cornutus*	*Z.scopas*
A. sparoides	82.8								
A .vaigiensis	1.1	80.0							
C. trifascialis			85.1						
C. weberi							1.1		
N. elegans				95.1					3.9
P. sulfureus					73.8	2.6			
P. diacanthus							86.8		
T. hardwicke								87.3	
Z. cornutus									89.0
Z. scopas									92.7
Part of *A. sparoides*	6.0								
Part of *A. vaigiensis*	1.0	9.1							
Part of *C. trifascialis*			2.6						
Part of *N. elegans*				1.6					
Part of *P. sulfureus*			1.1		4.3				
Part of *P. diacanthus*									
Part of *T. hardwicke*							2.6		

Part of *Z. cornutus*								6.2	
Part of *Z. scopas*									
Environment	8.0	10.9	9.5	2.2	20.9	7.9	9.2	4.6	2.8

Table S5. Performance of our final CNN model to identify 9 fish species

Raw model output was post-processed with following decision rule: outputs " part of species X" and "species X" are considered the same (i.e., the results of *A. sparoides* and *part of A. sparoides* are added together*); s*pecies are in columns with rows indicating the percentage of good identification for each species and only values over 1% are shown.

Species	*A. sparoides*	*A. vaigiensis*	*C. trifascialis*	*N.elegans*	*P.sulfureus*	*P.diacanthus*	*T.hardwicke*	*Z.cornutus*	*Z.scopas*
A. sparoides	89.0								
A. vaigiensis	2.1	89.1							
C. trifascialis			97.7						
C. weberi							1.1		
D. caruleus									3.9
N. elegans				95.7					
P. sulfureus			3.1		78.1	2.6			
P. diacanthus						86.8			
T. hardwicke						89.4			
Z.cornutus								95.2	
Z. scopas									92.7
Environment	8.0	10.9	9.5	2.1	20.9	7.9	9.2	4.6	2.8

SUPPLEMENTARY FIGURES

Supplementary Figure 1. Screenshot of the online application used for testing human performance in identifying fish on thumbnails. Picture of fish to identify is displayed on the left part. Name for species should be typed in the bottom text bar (with auto-completion). The help box with examples of the 9 species to identify is visible on the right.