UNCERTAINTY VISUALIZATION FOR NON-EXPERT USERS

Emma Beauxis-Aussalet, Lynda Hardman - CWI
{Emma, Lynda.Hardman}@cwi.nl

ICVSS 2014
Sicily ~ 13-19 July
International Computer Vision Summer School

Abstract

Computer vision uncertainty is not easily understood by end-users. It can yield trust issues impeding technology transfers, or misinterpretations of data, which can be critical. We propose visualization designs for end-user assessment of uncertainty. We investigate the requirements of non-experts, and the corresponding visualizations. Our study primarily concerns an application for scientific research. Our insights are generalizable to other domains.

		Classification from Ground-Truth				Basic Metrics				Advanced Metrics					
		Anchovy	Barracuda	Clown Fish	Other	TP	FP	FN	TN	Precision TP/(TP+FP)	Recall or TP Rate TP/(TP+FN)	FP Rate FP/(FP+TN)	Accuracy (TP+TN)/All	F1 score 2TP/(2TP+FP+FN)	Matthews Cor. Coefficient
Classification from Computer Vision	Anchovy	85	12	0	15	85	27	15	273	0.76	0.85	0.09	0.90	0.80	0.73
	Barracuda	12	75	4	9	75	25	25	275	0.75	0.75	0.08	0.88	0.75	0.67
	Clown Fish	0	9	95	6	95	15	5	285	0.86	0.95	0.05	0.95	0.90	0.87
	Other	3	4	1	70	70	8	30	292	0.90	0.70	0.03	0.91	0.79	0.74

User Needs

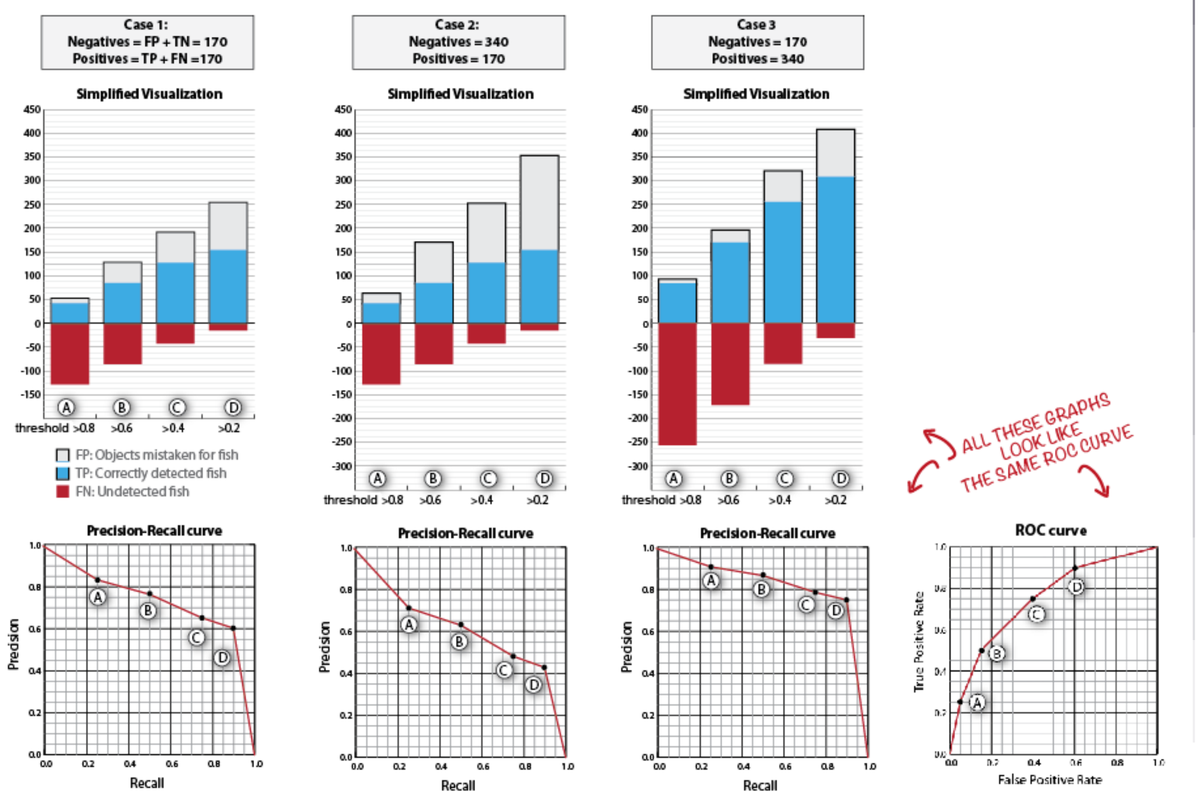
We iteratively interviewed users along the design stages. They need to visualize classification errors for 1) all classes; 2) the classes often confused with one another; and 3) for uncertainty factors such as image quality. We present our latest visualization design. Future work will evaluate their usability with controlled experiments, and open-ended data analysis tasks with the Fish4Knowledge dataset.

Issues with Confusion Matrix

Inter-classes confusions are complex: FN of one class are FP for another. Errors can be compared over parameter settings, or uncertainty factors (e.g., image quality). Advanced metrics (e.g., Precision/Recall, F1 score) are complicated for non-experts. They may not know which metrics suit their use case, or misinterpret them. E.g., high TN may conceal uncertainty by yielding low *FP Rate* and high *Accuracy*.

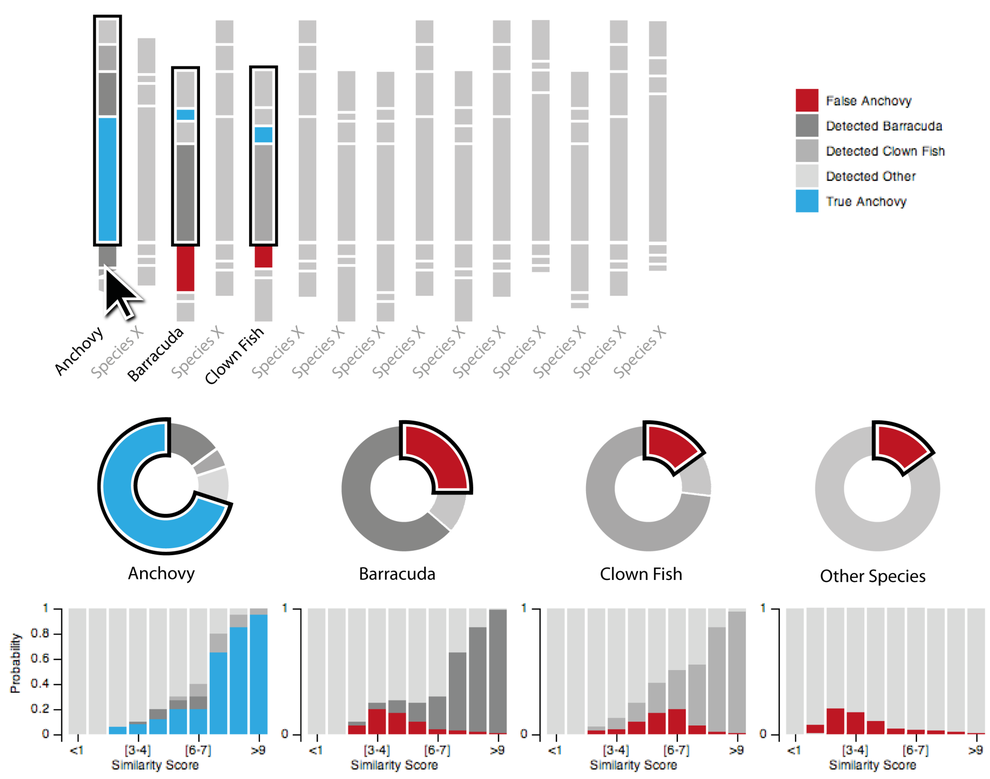
Choice of Metrics

To ease non-experts' interpretation, we display only TP, FN and FP. Error evaluation is more tangible, as well as the impact of parameter settings. It suits most use cases, whereas inappropriate expert visualizations (e.g., ROC curve) may conceal relevant errors.



Detail of Confusions

Inter-classes confusion is detailed for each class. We select the 2 classes producing the most FP and FN. We show the proportions of misclassified items for these classes, as well as the average errors for other classes. Error distributions over *similarity scores* indicate how much classes look alike.



Future Work with Logistic Regression

We are experimenting with logistic regression to reduce uncertainty. Fitted with *similarity scores*, it provides classification probabilities for each fish. Fish counts can be improved with these probabilities (e.g., a fish is counted as 0.8 Anchovy and 0.2 Barracuda). Our visualization can show the benefits to end-users. We will investigate how visualization can support i) the understanding of uncertainty before and after logistic regression, ii) user trust in computer vision and logistic regression.

References

[1] E. Beauxis-Aussalet, L. Hardman., Uncertainty-Aware Visualization of Fish Populations, *Dema at AVI*, 2014
[2] E. Beauxis-Aussalet, E. Arslanova, L. Hardman, A Case Study of Trust Issues in Scientific Video Collections, *MAED Workshop at ACM MM*, 2013
[3] E. Beauxis-Aussalet, et al., A Video Processing and Data Retrieval Framework for Fish Population Monitoring, *MAED Workshop at ACM MM*, 2013
[4] Fish4Knowledge project - UI Prototype: `http://f4k.project.cwi.nl` - Video: `https://www.youtube.com/watch?v=AFV-FiKUFyI`