How to increase accuracy of crowdsourcing campaigns?

Nurmukhametov Oleg [1] , Baklanov Artem [2], Fritz Steffen [3], Khachay Mikhail [1], Salk Carl [3],See Linda [3], Shchepashchenko Dmitry [3] .

1 — N.N. Krasovskii Institute of Mathematics and Mechanics (Russian Academy of Sciences);
2 — Advanced System Analysis Program, IIASA;
3 — Ecosystems Services & Management Program, IIASA.

Challenge

How to aggregate votes from non-experts?

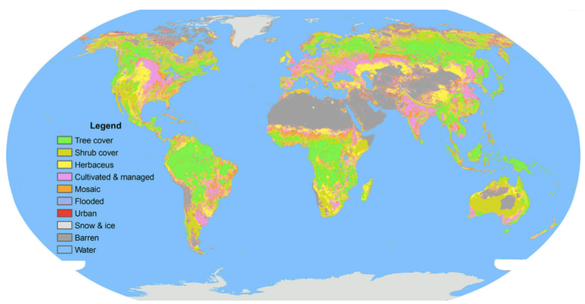
"Cropland Capture" game

iOS

Over **5 millions** opinions from non-experts

HOW?

Land cover map



Expert quality decisions about 190 000 images

Results

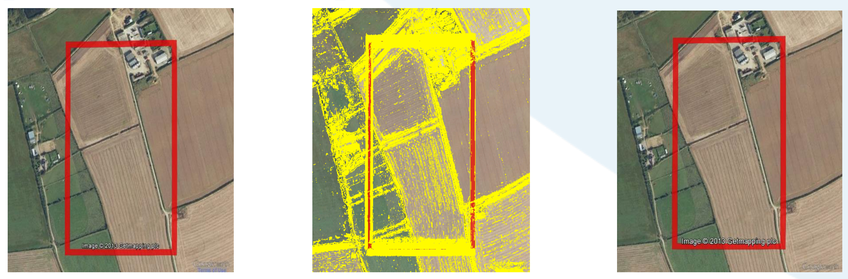
We increased accuracy of "Cropland Capture" data from **76% to 87%**

- ✓ We proposed and tested the two-step procedure for generic crowdsourcing campaigns to reduce noise and to increase an efficiency of a task allocation;
- ✓ We improved estimate of simple heuristics (majority voting);
- ✓ We proposed ways to aggregate votes which significantly outperform heuristic rules.

Approach

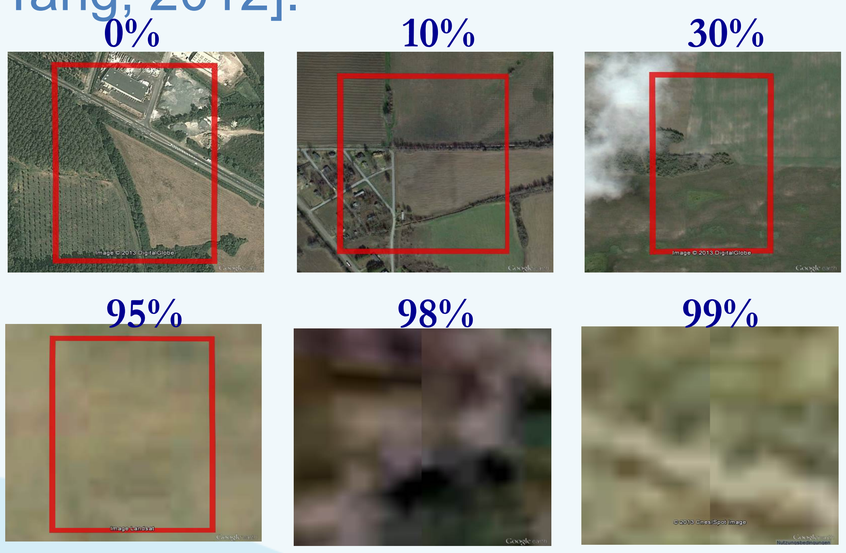
Analysis of images

1) Detection of similar images using pHash (perceptual hash) [Zauner, 2010].



➔5% of images are not unique

2) Detection of low quality images using Blur detection algorithm [H Tang, 2012].



➔2% of images are discarded

Analysis of votes

Volunteers change opinions.

We compared simple heuristic rules for aggregation votes on individual level.

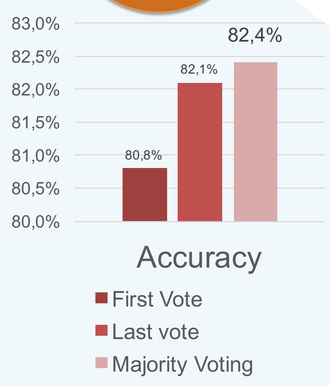
Heuristic	Assumption
1st Vote	*Volunteers lose attention.* Performance of volunteer is decreasing↓.
Last Vote	*Volunteers learn over time.* Performance of volunteer is increasing↑.
Majority Voting	*The Wisdom of Crowds.* Combined opinion is a right answer.

Heuristic reduces dimensionality

5 millions votes → 2.7 million

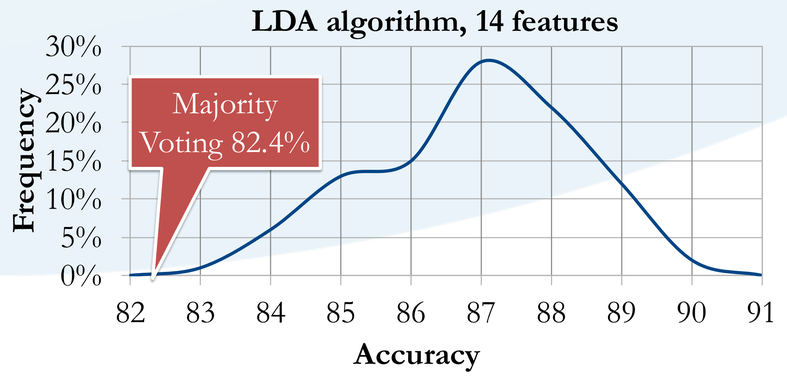
Disagreement on dataset

	1st Vote	Last Vote	Majority vote
1st Vote	-	4%	3%
Last Vote	4%	-	2%
Majority Vote	3%	2%	-



Decision-making

We applied state of the art machine learning algorithms to obtain the best way for aggregation of votes on the expert validated dataset and then predict expert's decision for any image using voting protocol.



Plot depicts an accuracy of the best suited machine learning algorithm (**linear discriminant analysis)** for 100 random splits (60/40) of the experts dataset.

Business cards and sweets=)