

Supplementary Information

Crowd-driven deep learning tracks Amazon deforestation

¹*McCallum, Ian*

²*Walker, Jon*

¹*Fritz, Steffen*

³*Grau, Markus*

²*Hannan, Cassie*

²*Hsieh, I-Sah*

²*Lape, Deanna*

²*Mahone, Jen*

²*McLester, Caroline*

²*Mellgren, Steve*

²*Piland, Nolan*

¹*See, Linda*

⁴*Svolba, Gerhard*

⁵*de Villiers, Murray*

¹ International Institute for Applied Systems Analysis, Laxenburg, Austria

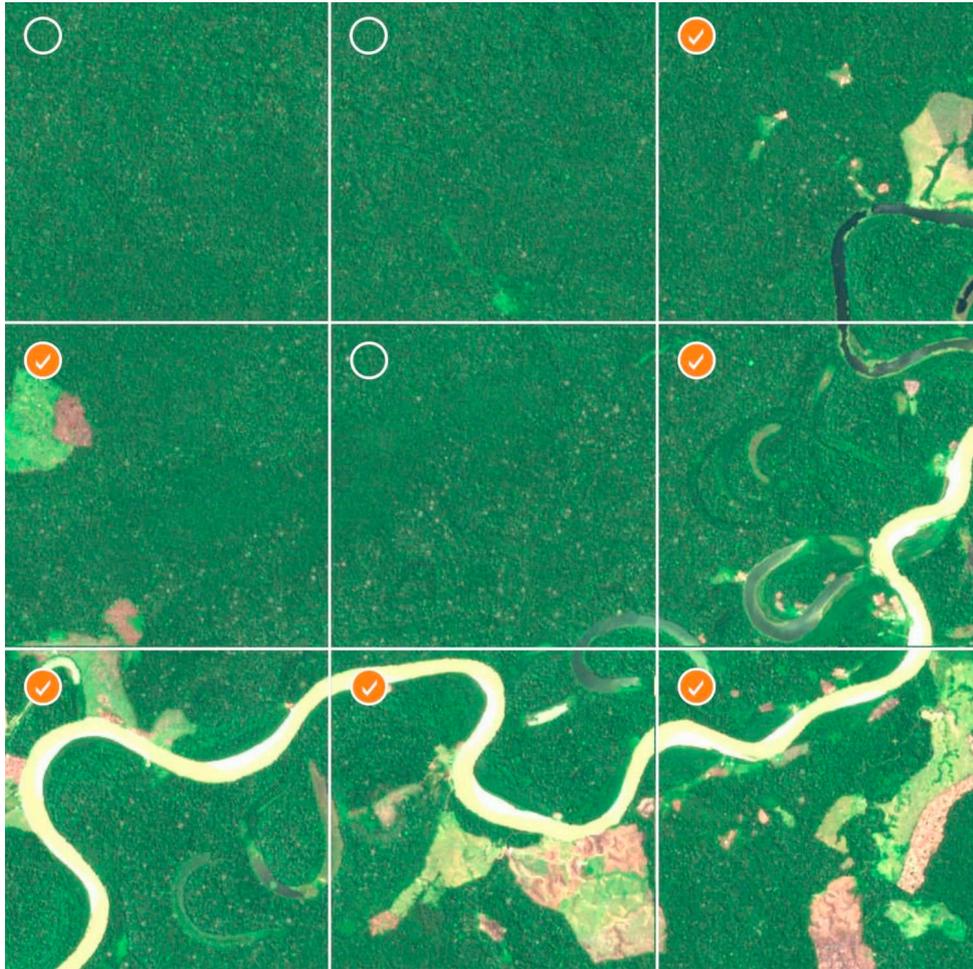
² SAS Campus Drive, Cary, NC 27513, United States

³ SAS Institute AG, Richtistrasse 11, 8304 Wallisellen, Switzerland

⁴ SAS Austria, Trabrennstrasse 2B, 1020 Wien, Austria

⁵ SAS Netherlands B.V. Flevolaan 69, 1272 PC Huizen, Netherlands

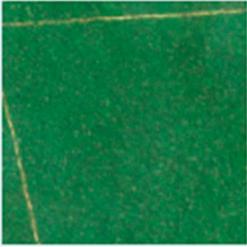
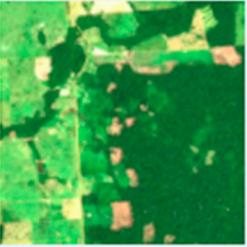
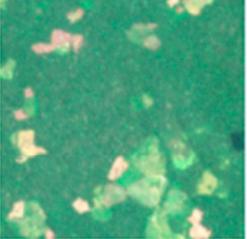
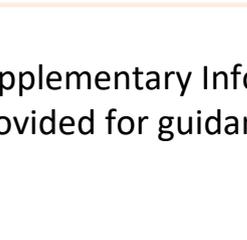
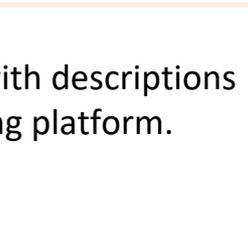
Results



Supplementary Information Figure S1. The crowdsourcing user interface which prompts users to check any images they deem to contain deforestation in a 3x3 window. Unchecked boxes are then assumed to contain no deforestation.

Examples



Human Impact	Natural Deforestation	Comparisons	
			▲
Road	Roads connecting fields	Settlement	
			
Roads and fields	Clear cuts or existing fields (sun and satellite sensor angle create bright reflectance)	Intentional clearings along a river	
			
Roads and clearings	Road through natural landscape	Intentional clearings	▼
			
Roads and intentional clearings		Roads and intentional clearings	

Supplementary Information Figure S2. Example images with descriptions provided for guidance on the deforestation crowdsourcing platform.

Validation of the Crowd

Supplementary Table S1. Spatial accuracy between the crowd and PRODES deforestation product.

Confusion Matrix and Statistics

```

      Reference
Prediction 0  1
0      822  38
1      48  155

      Accuracy : 0.9191
      95% CI   : (0.901, 0.9348)
No Information Rate : 0.8184
P-Value [Acc > NIR] : <2e-16

      Kappa : 0.7332

McNemar's Test P-Value : 0.3318

      Sensitivity : 0.9448
      Specificity : 0.8031
      Pos Pred Value : 0.9558
      Neg Pred Value : 0.7635
      Prevalence : 0.8184
      Detection Rate : 0.7733
      Detection Prevalence : 0.8090
      Balanced Accuracy : 0.8740

      'Positive' Class : 0
```

Supplementary Table S2. Spatial accuracy between the crowd and global annual tree cover loss product.

```
> (mat2)
```

Confusion Matrix and Statistics

```

      Reference
Prediction 0  1
0      845  15
1     107  96

      Accuracy : 0.8852
      95% CI   : (0.8645, 0.9038)
No Information Rate : 0.8956
P-Value [Acc > NIR] : 0.8748

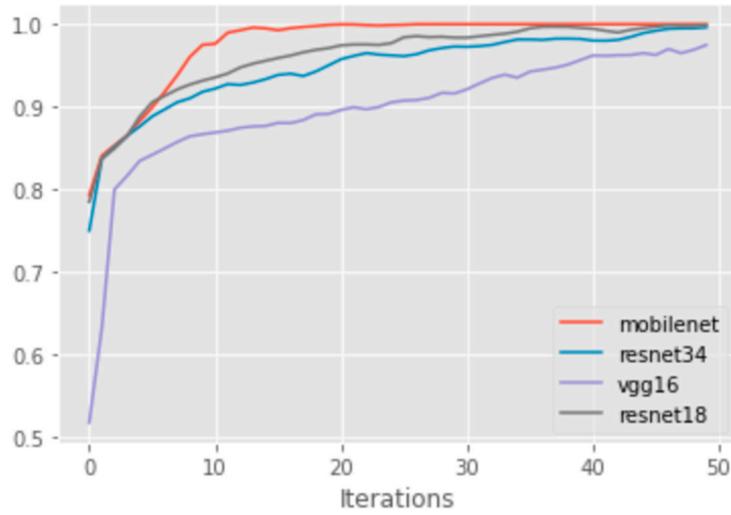
      Kappa : 0.5508

McNemar's Test P-Value : <2e-16

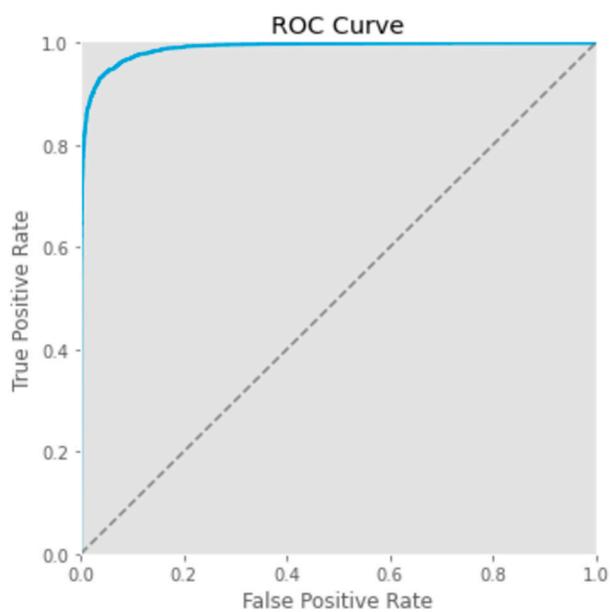
      Sensitivity : 0.8876
      Specificity : 0.8649
      Pos Pred Value : 0.9826
      Neg Pred Value : 0.4729
      Prevalence : 0.8956
      Detection Rate : 0.7949
      Detection Prevalence : 0.8090
      Balanced Accuracy : 0.8762
```

'Positive' Class : 0

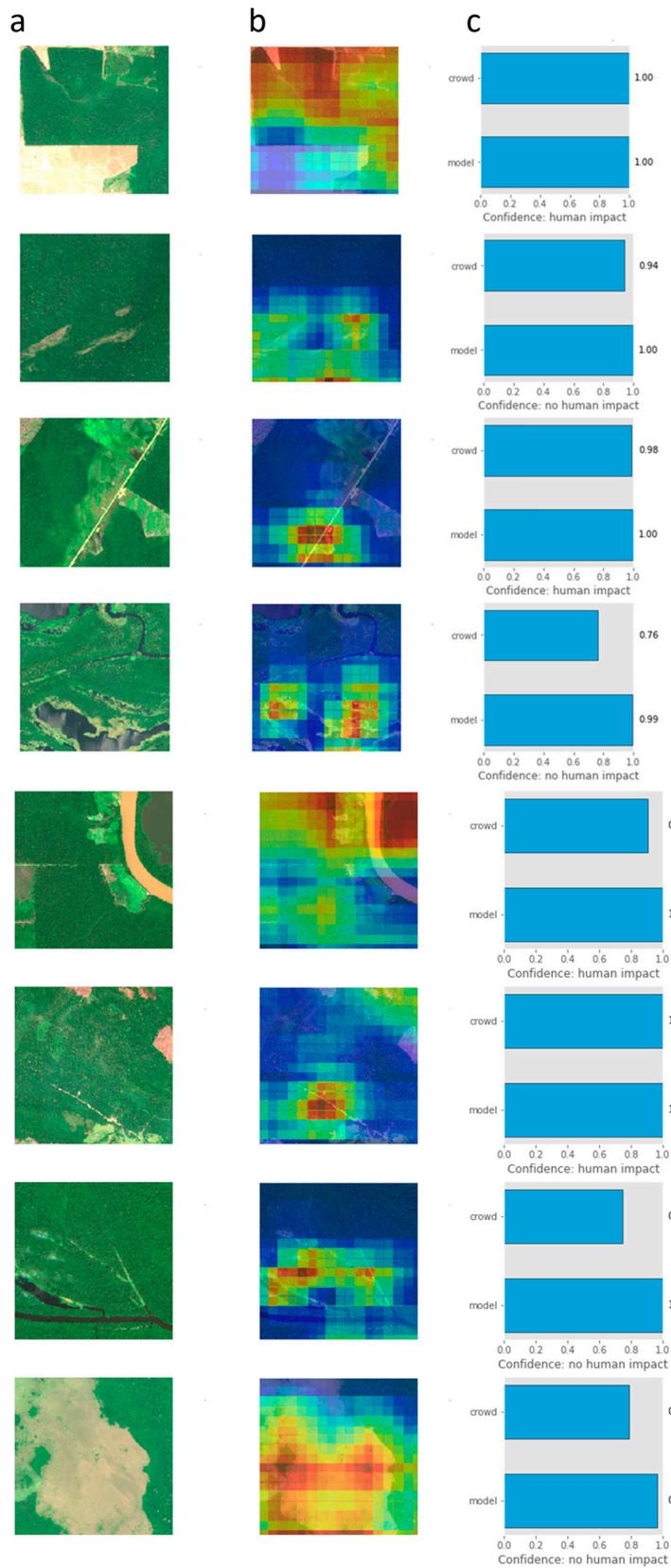
Deep Learning



Supplementary Information Figure S3. The mean training-set accuracy achieved by each model architecture computed over five separate trials.

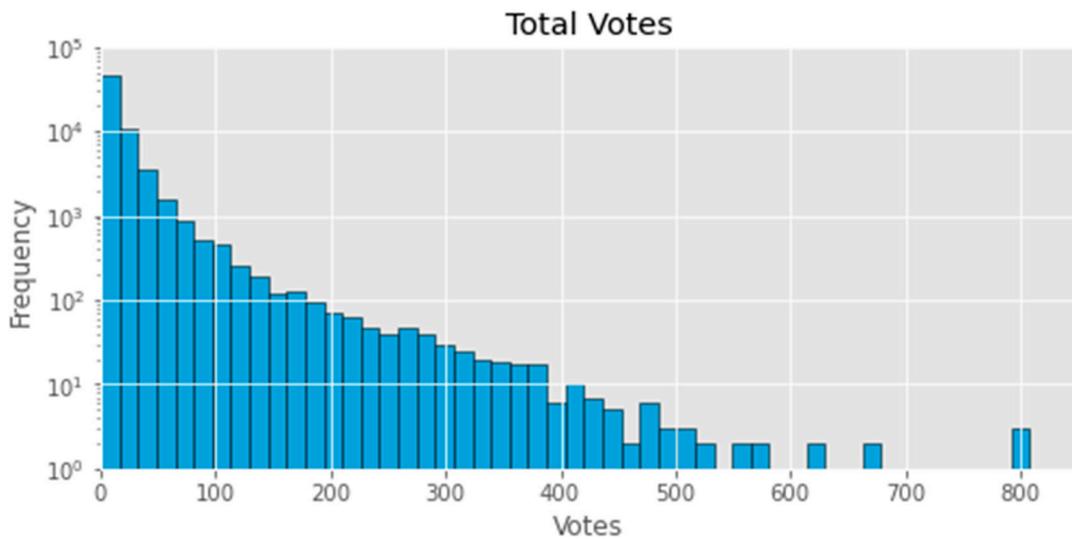


Supplementary Information Figure S4. The Receiver Operating Characteristic curve illustrating the change in true positive and false positive classifications as the threshold for a positive class label is varied. Optimal threshold of 0.39 as determined using Youden's Index is marked in black.



Supplementary Information Figure S5. Examples of deep learning model results for a) eight test dataset images from across the Brazilian Amazon representing signs of deforestation activity and no activity (including natural breaks in the

canopy), b) the resulting activation layers from the model showing which part of the image triggered the classifier (warmer colours imply model activation) and c) the resulting confidence of the crowd and model for deforestation or no deforestation.



Supplementary Information Figure S6. Frequency with which images require more than six votes before being successfully labeled. Frequency is on a log scale.

Supplementary Table S3. Accuracy of the crowd on a random sample of 200 images.

		Crowd Label		
		Human Impact	No Human Impact	Undecided
Expert Label	Human impact	68	15	5
	No Human Impact	2	109	1

Supplementary Table S4. Accuracy of the ResNet18 model on test images.

		ResNet18 Label	
		Human Impact	No Human Impact
Crowd Label	Human impact	3282	273
	No Human Impact	176	5043

Supplementary Table S5. Spatial accuracy of the crowd-driven AI model prediction compared with global annual tree cover loss over the Peruvian Amazon.

Confusion Matrix and Statistics

	Reference	
Prediction	0	1
0	0	0
1	3	16

Accuracy : 0.8421
95% CI : (0.6042, 0.9662)
No Information Rate : 0.8421
P-Value [Acc > NIR] : 0.6478

Kappa : 0

Mcnemar's Test P-Value : 0.2482

Sensitivity : 0.0000
Specificity : 1.0000
Pos Pred Value : NaN
Neg Pred Value : 0.8421
Prevalence : 0.1579
Detection Rate : 0.0000
Detection Prevalence : 0.0000
Balanced Accuracy : 0.5000

'Positive' Class : 0