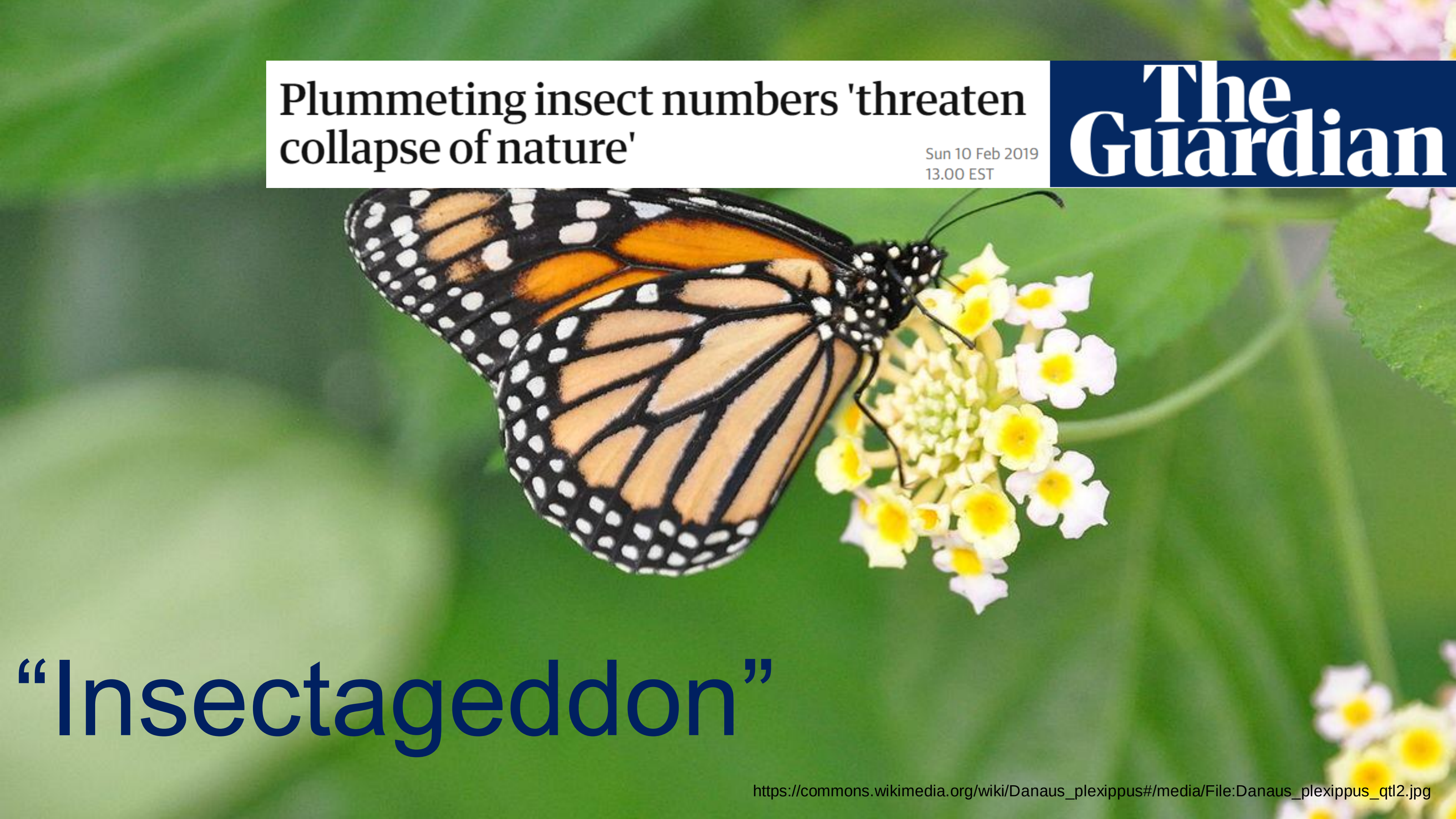
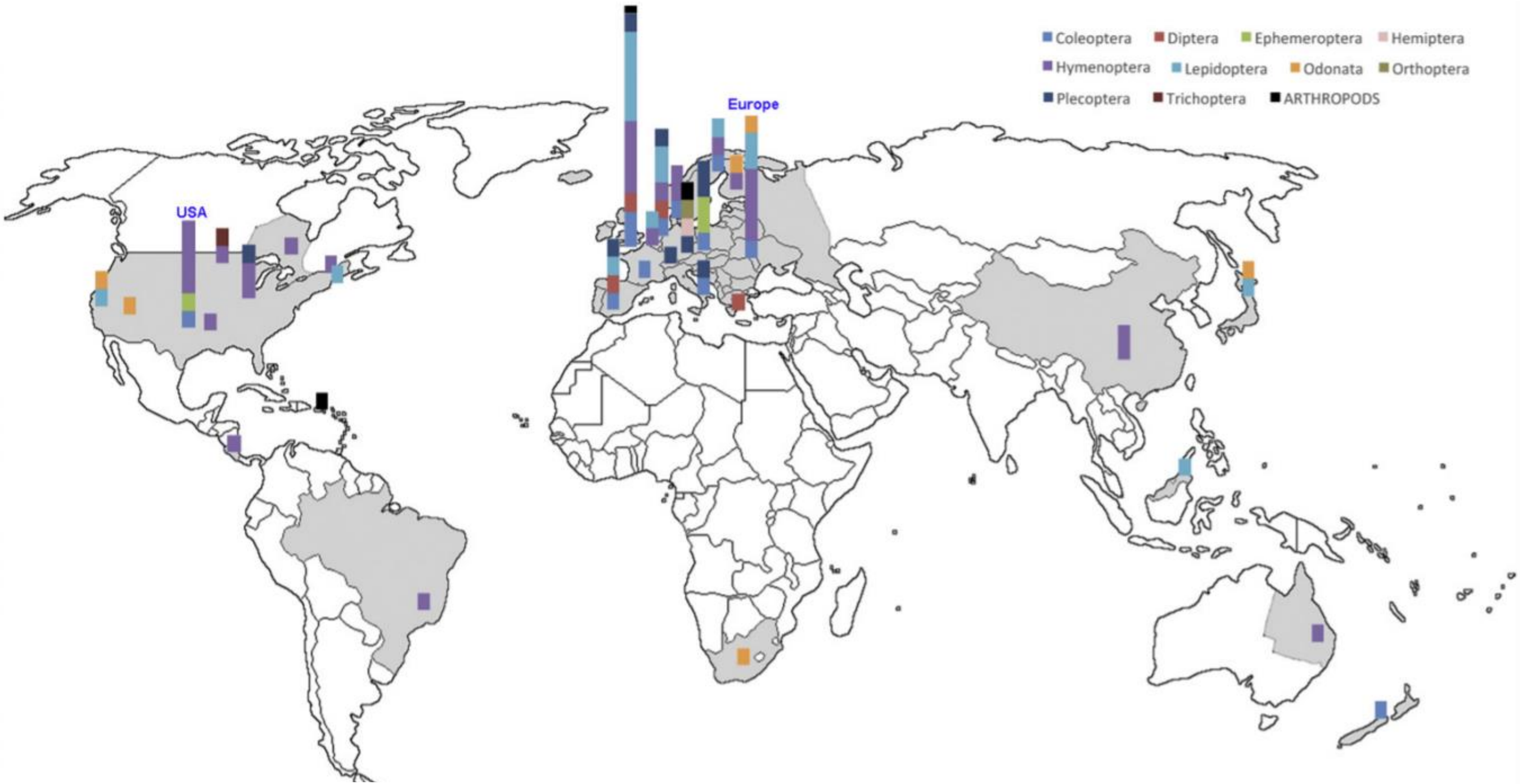
Fine-Grained Recognition With Humans-in-the-Loop

Oisin Mac Aodha

University of Edinburgh

www.oisin.info


Plummeting insect numbers 'threaten collapse of nature'
Sun 10 Feb 2019
13.00 EST
The Guardian
“Insectageddon”
https://commons.wikimedia.org/wiki/Danaus_plexippus#/media/File:Danaus_plexippus_qtl2.jpg



Review

Worldwide decline of the entomofauna: A review of its drivers

Francisco Sánchez-Bayo [a], Kris A.G. Wyckhuys [b, c, d]

How can we automate wildlife monitoring on a global scale?

iNaturalist 2017

5,089 classes
Bounding Boxes

iNaturalist 2018

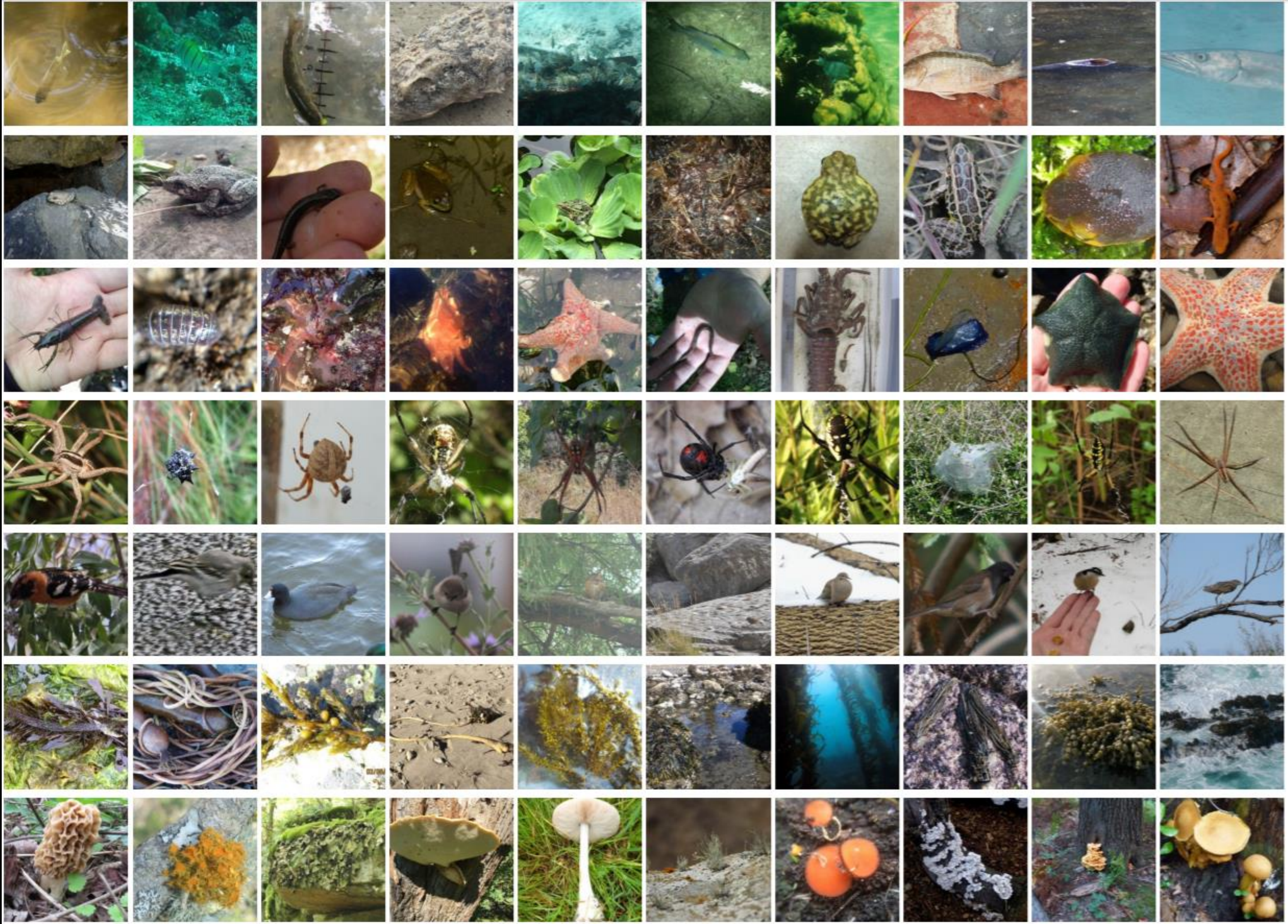
8,142 classes
Taxonomy

iNaturalist 2019

1,100 “hard” classes

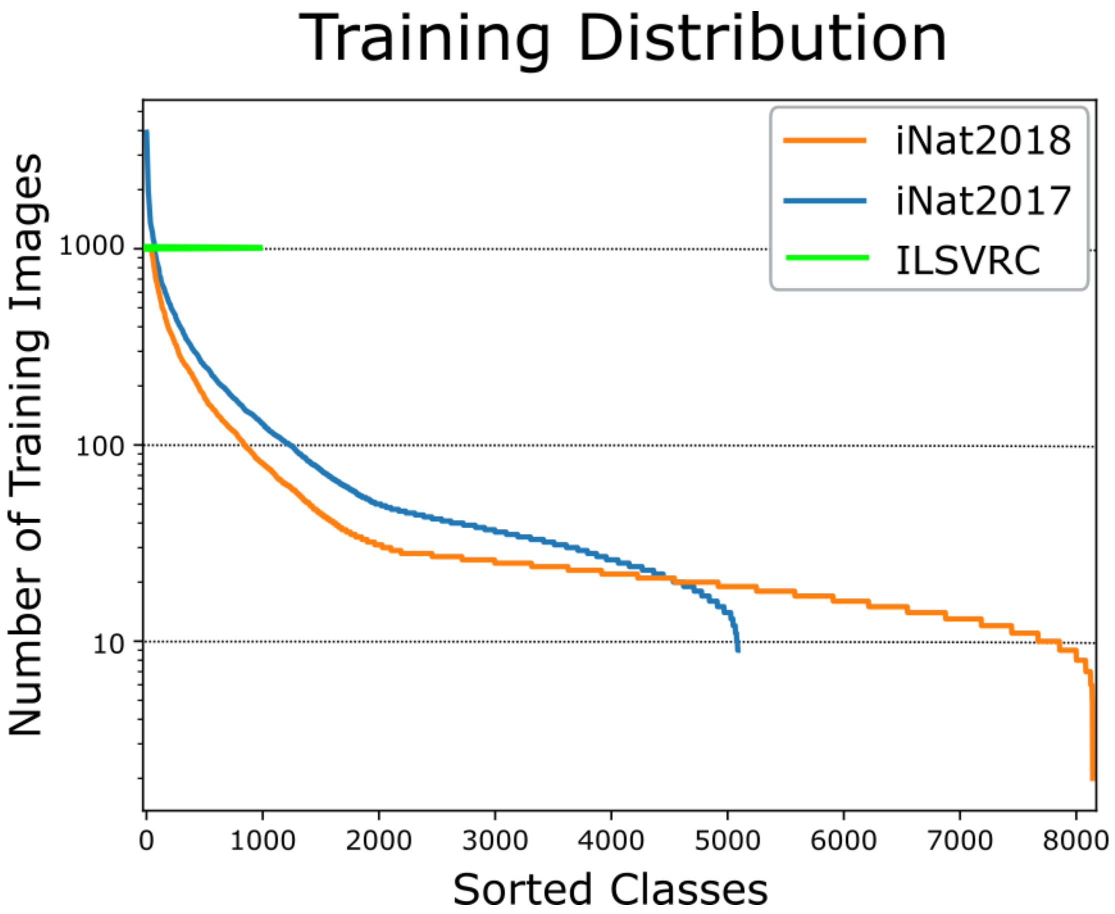
The iNaturalist Species Classification and Detection Dataset **CVPR 2018**
Van Horn, **Mac Aodha**, Song, Cui, Sun, Shepard, Adam, Perona, Belongie

iNaturalist.org

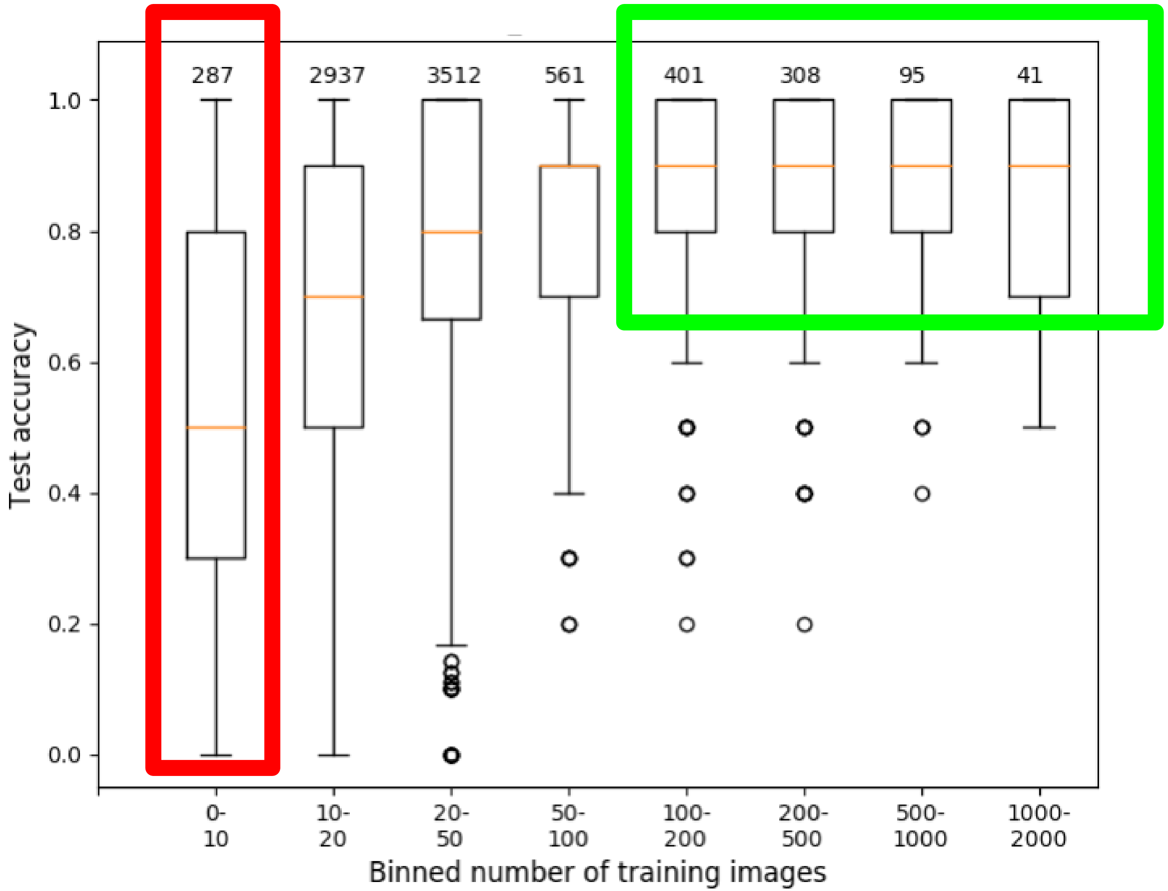


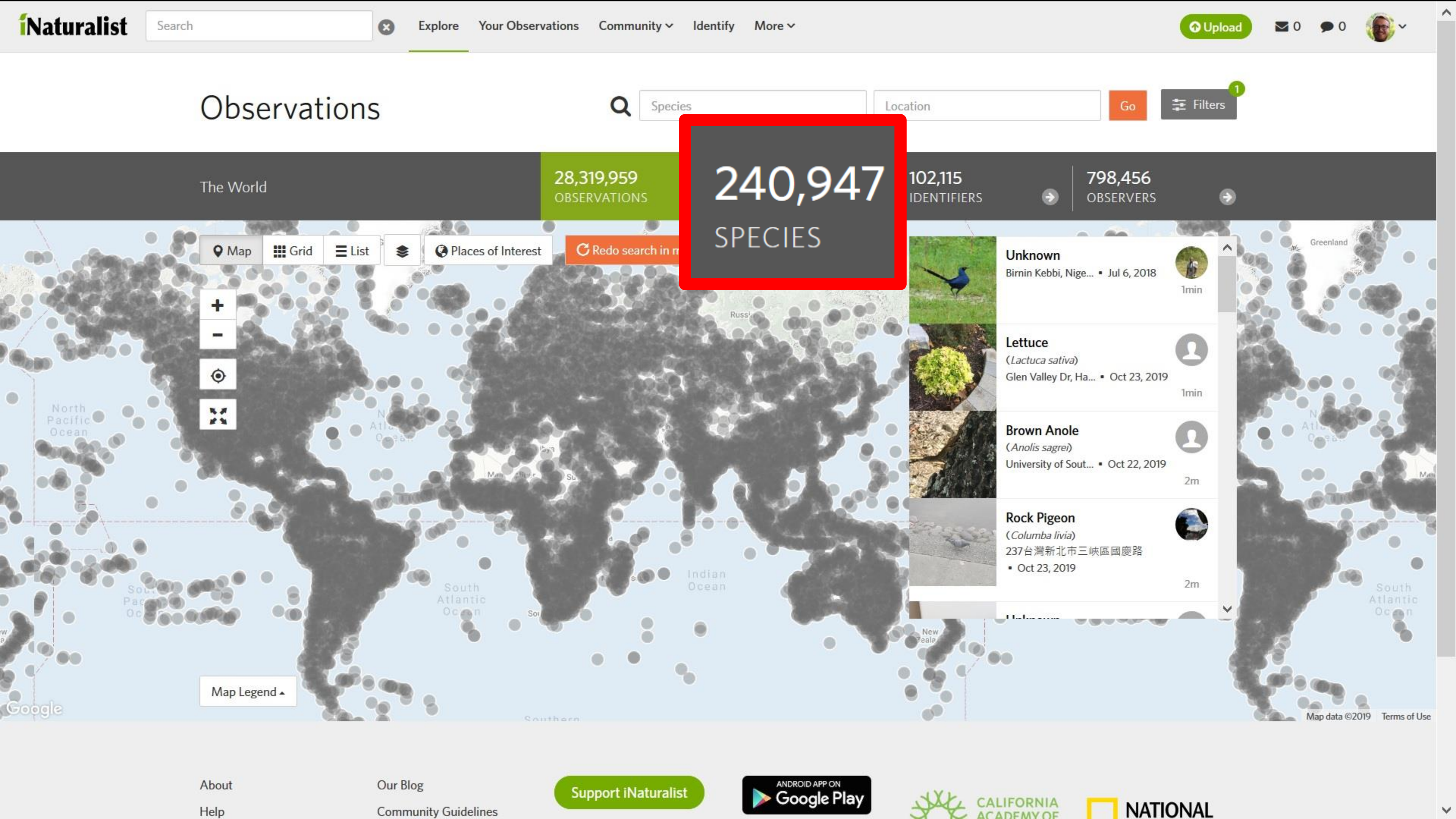
Ardea cinerea

Ardea cocoi

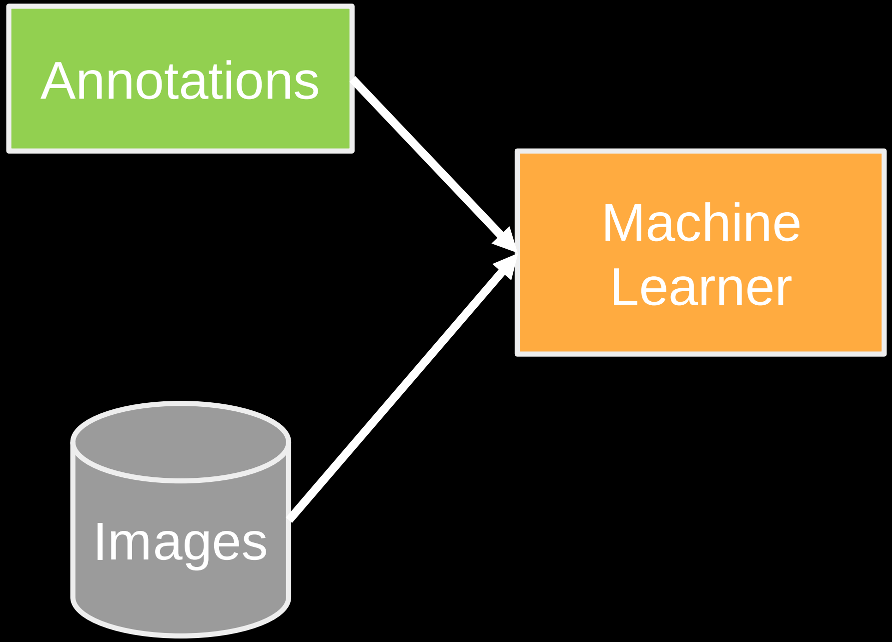

Training Distribution
iNat2018
iNat2017
ILSVRC
Number of Training Images
1000
100
10
0
1000
2000
3000
4000
5000
6000
7000
8000
Sorted Classes

iNaturalist 2018 – Winner's Top 1 Accuracy

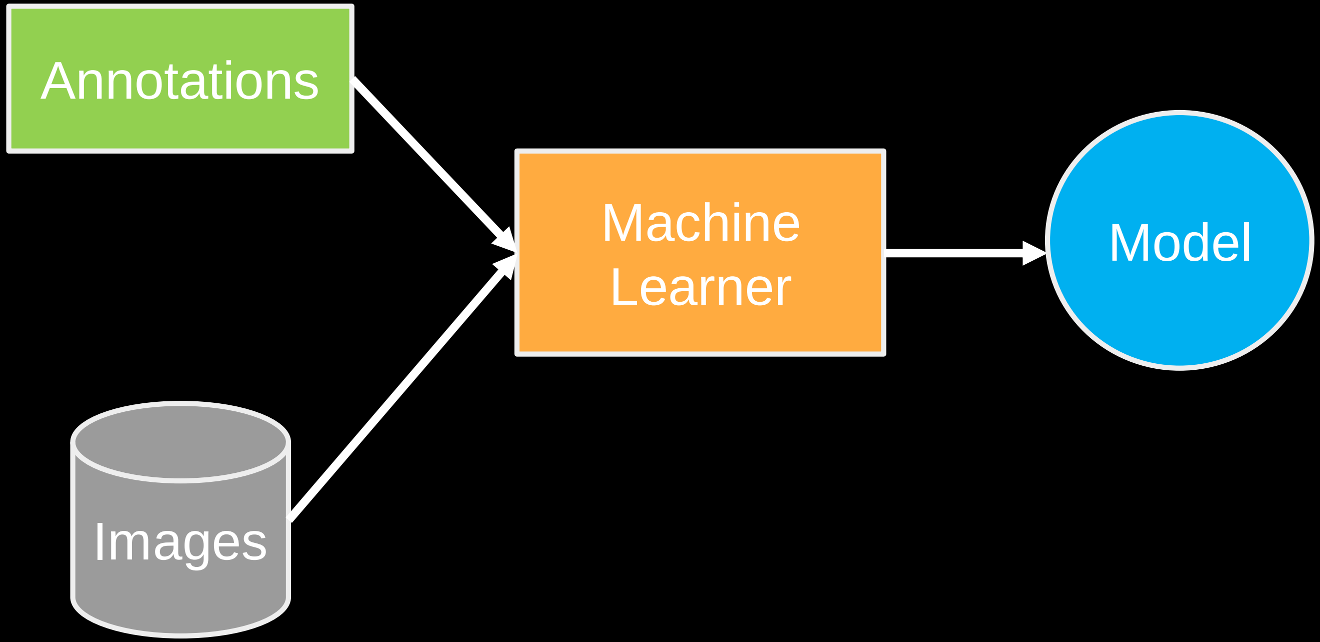



iNaturalist
Search
Explore
Your Observations
Community
Identify
More
Upload
Observations
Species
Location
Go
Filters
The World
28,319,959
OBSERVATIONS
240,947
SPECIES
102,115
IDENTIFIERS
798,456
OBSERVERS
Map
Grid
List
Places of Interest
Unknown
Birnin Kebbi, Nige... • Jul 6, 2018
1min
Lettuce
(Lactuca sativa)
Glen Valley Dr, Ha... • Oct 23, 2019
1min
Brown Anole
(Anolis sagrei)
University of Sout... • Oct 22, 2019
2m
Rock Pigeon
(Columba livia)
237台灣新北市三峽區國慶路
• Oct 23, 2019
2m
North Pacific Ocean
South Atlantic Ocean
Indian Ocean
Greenland
Map Legend
Google
Map data ©2019
Terms of Use
About
Help
Our Blog
Community Guidelines
Support iNaturalist
ANDROID APP ON
Google Play
CALIFORNIA ACADEMY OF
NATIONAL

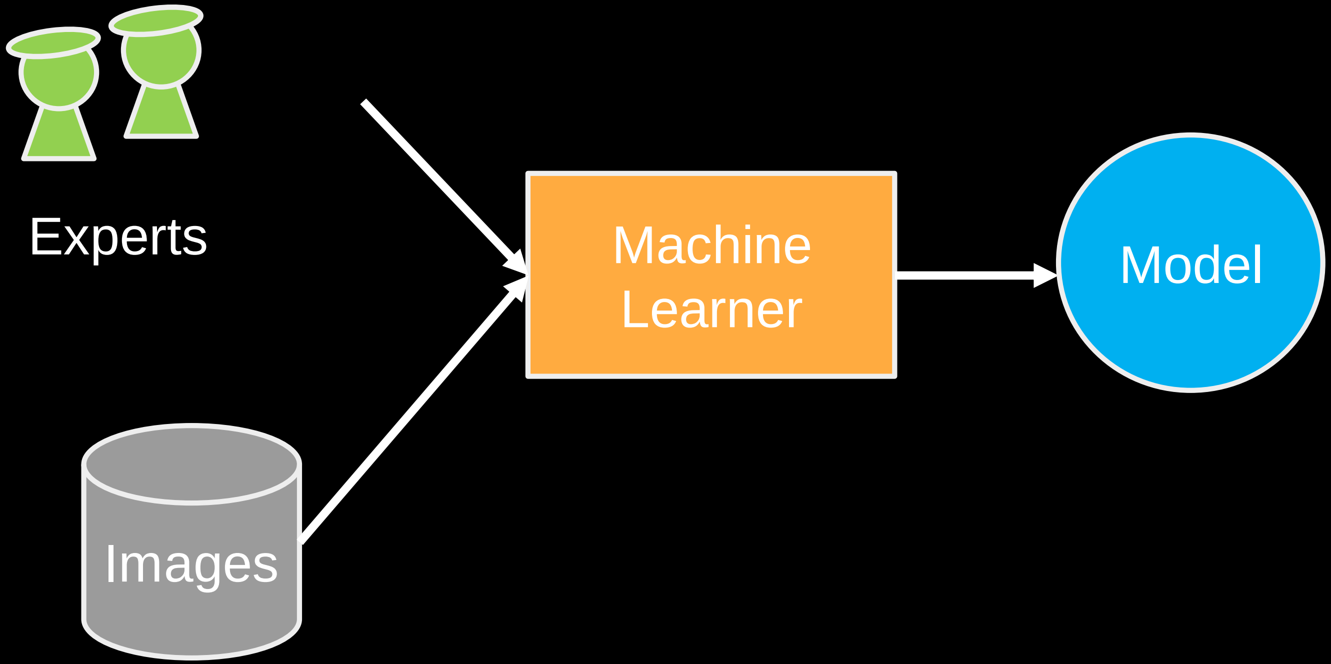
Conventional Machine Learning Pipeline



Conventional Machine Learning Pipeline

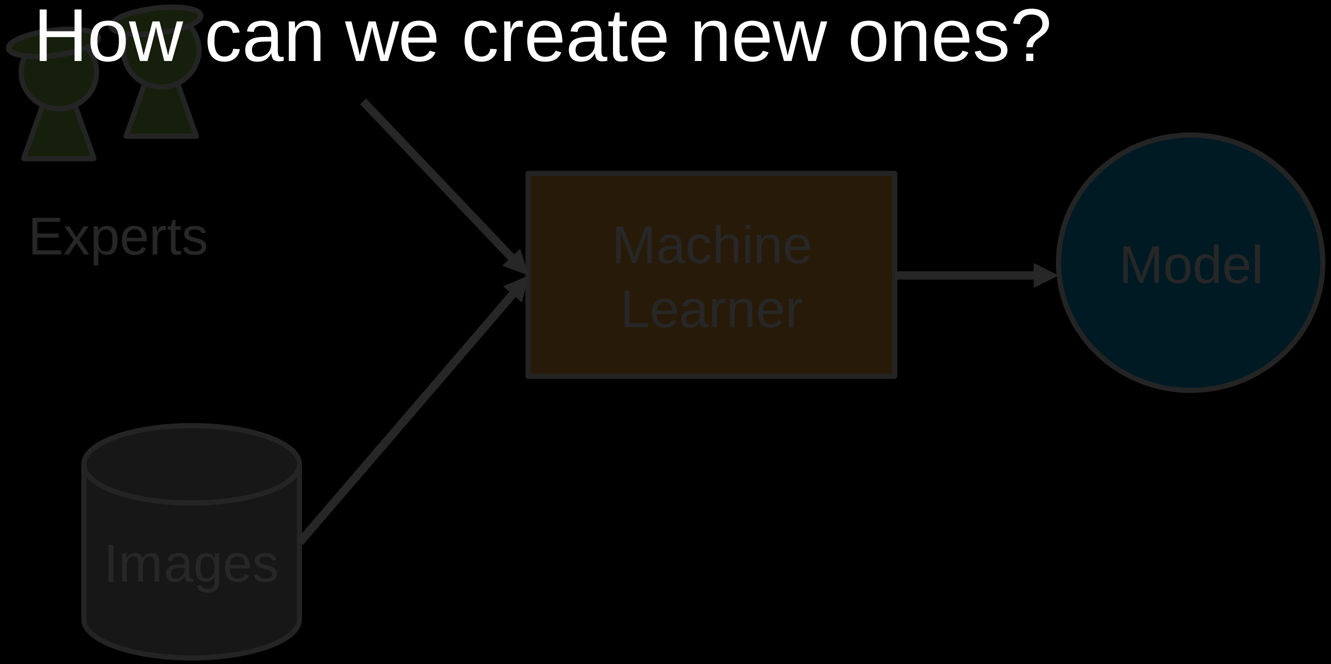


Conventional Machine Learning Pipeline

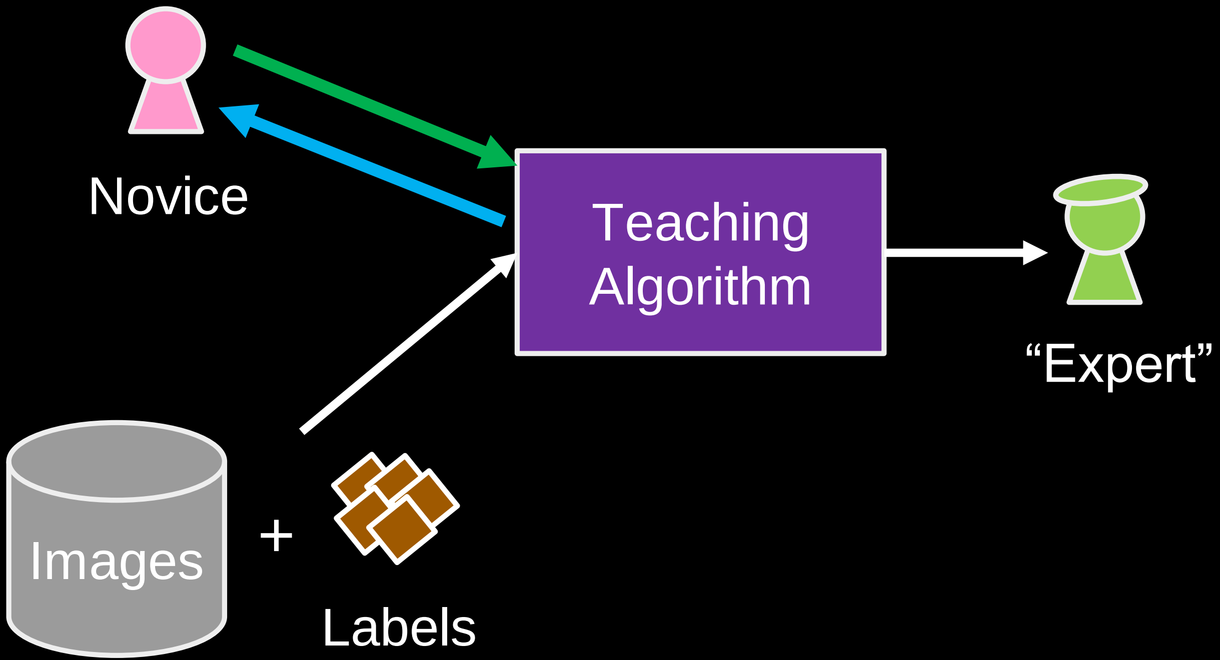


Problem: Limited number of experts.

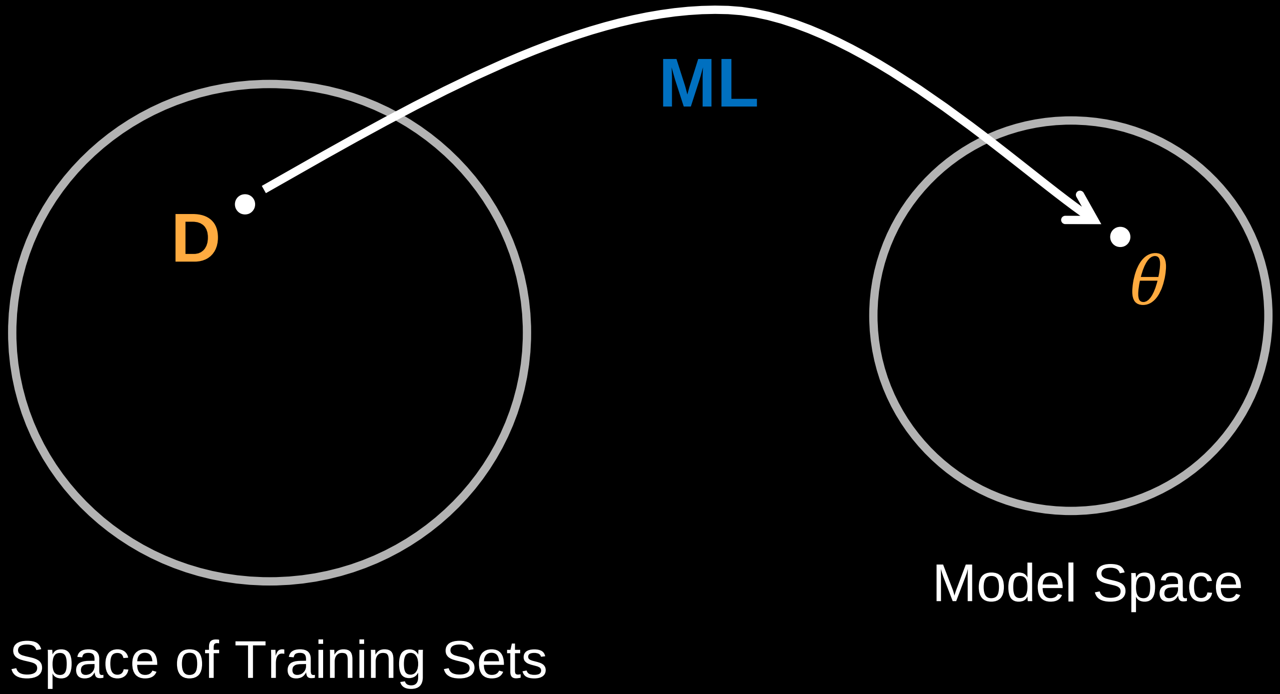
How can we create new ones?



Machine Teaching

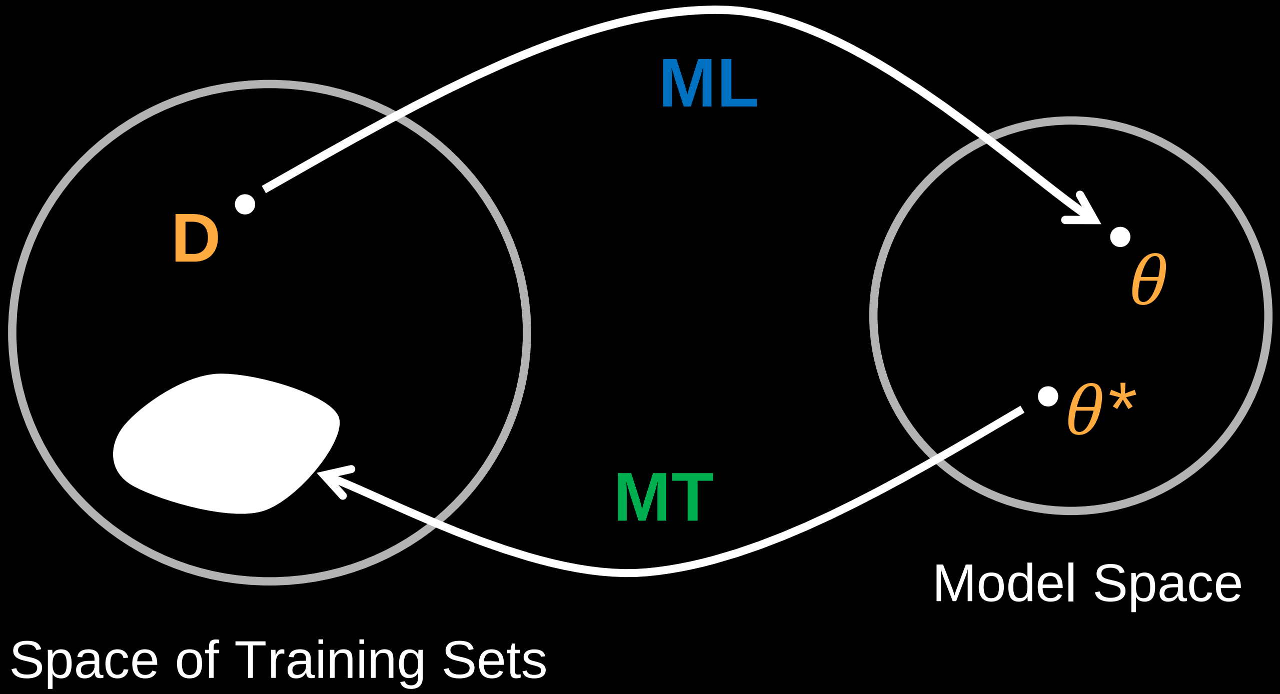


Machine Learning vs Machine Teaching



Zhu AAAI'15

Machine Learning vs Machine Teaching



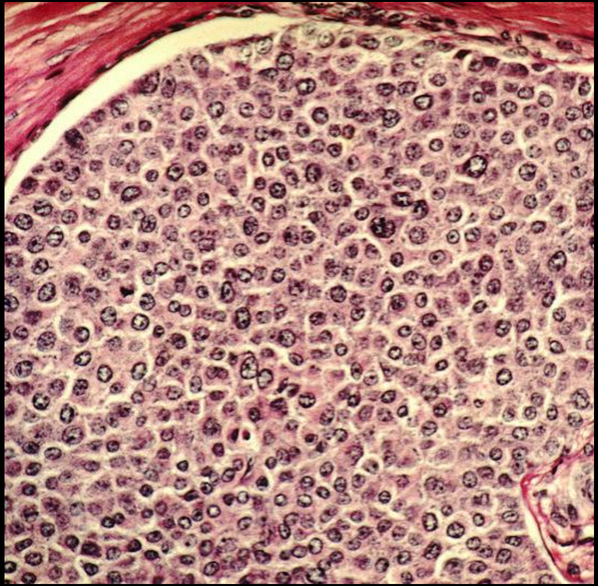
Zhu AAAI'15

Can we design teaching algorithms that enable humans to become better at fine-grained categorization?

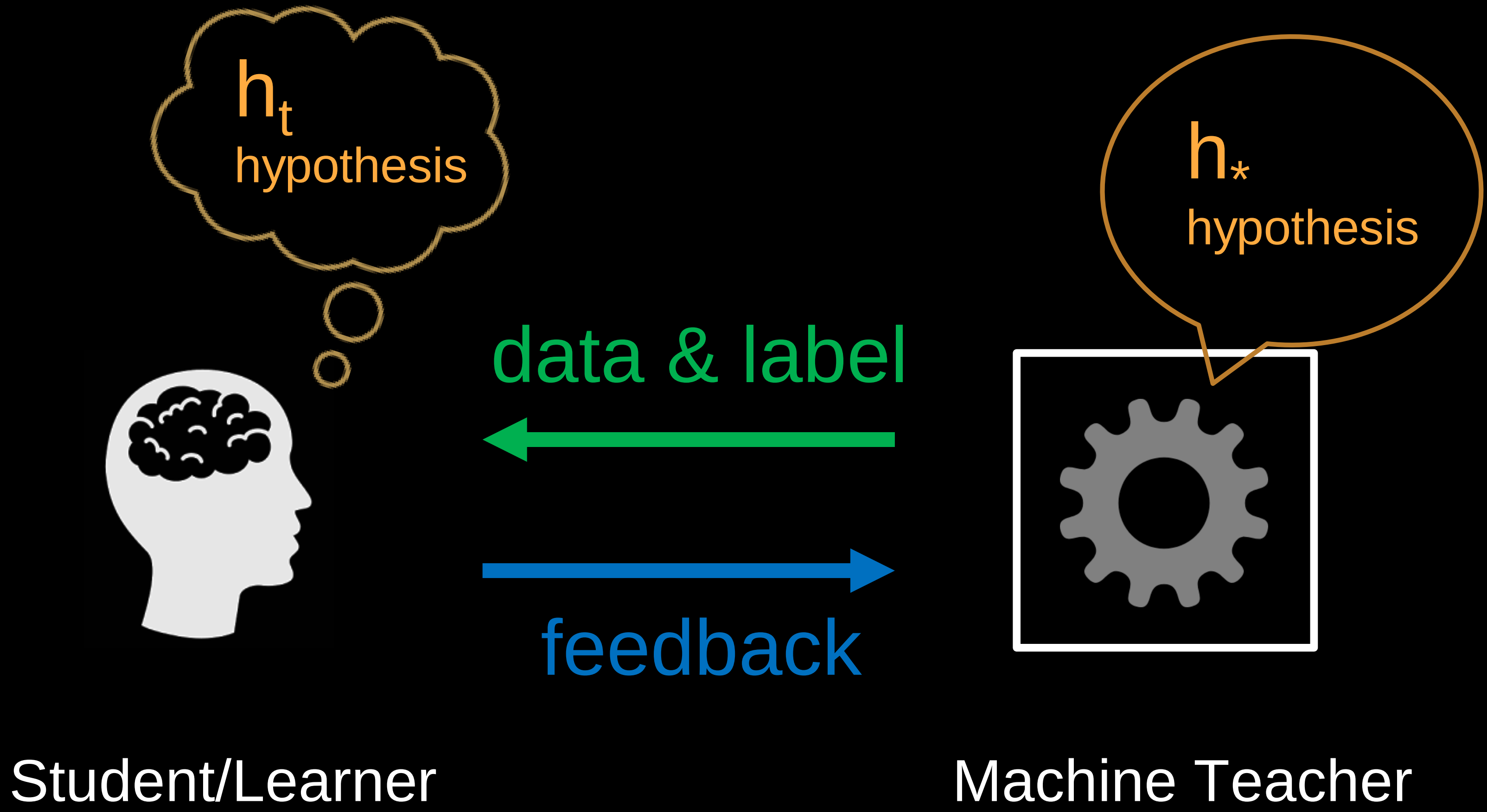
Why Visual Expertise?

Poisonous?

Cancer?



Forgery?


h_t
hypothesis
h_*
hypothesis
data & label
feedback
Student/Learner
Machine Teacher

Teaching Visual Expertise

⋮

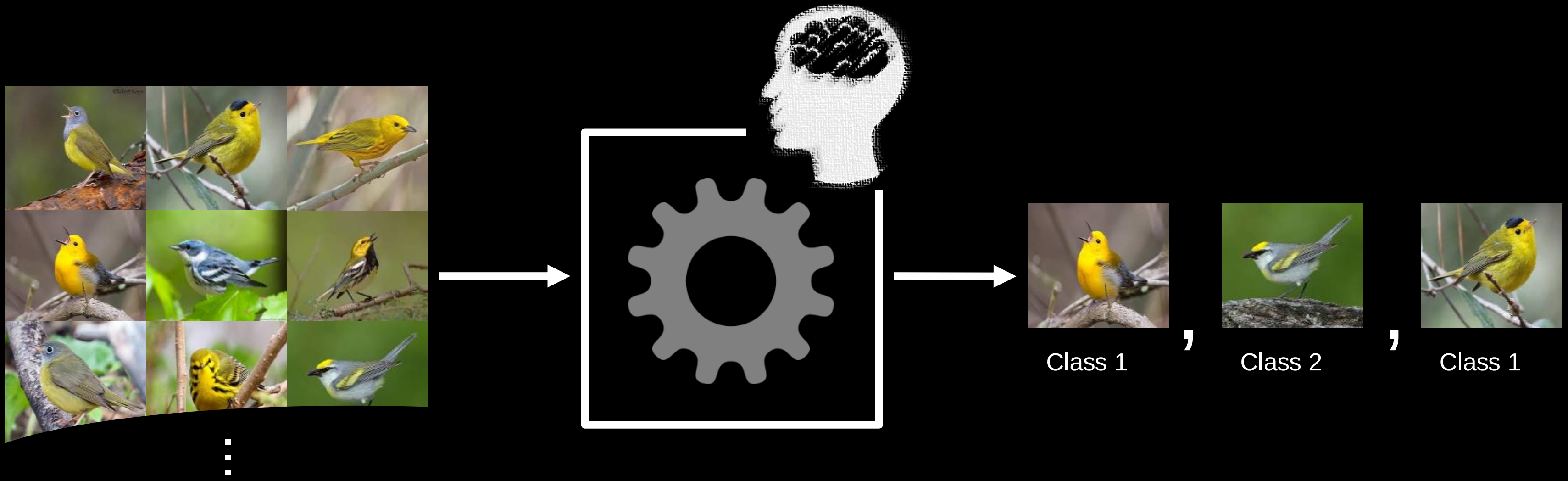
Set of images with
class labels

Teaching Visual Expertise

Set of images with class labels

Teaching algorithm & student model

Teaching Visual Expertise



Set of images with class labels

Teaching algorithm & student model

Sequence of teaching images

Machine Teaching Landscape

Theoretical
Goldman & Kearns 1995
Zhu 2013
Chen et al. 2018
...

Spaced Repetition
Leitner 1972
Settles & Meeder 2016
Hunziker et al. 2019
...

Decision Making
Bak et al. 2016
...

Visual Categories
Singla et al. 2014
Johns et al. 2015
Chen et al. 2018
...

Connecticut Warbler
or MacGillivray's Warbler

https://www.inaturalist.org/observations/9869215

Connecticut Warbler
~~or MacGillivray's Warbler~~

https://www.inaturalist.org/observations/9869215

Connecticut Warbler

MacGillivray's Warbler

https://www.inaturalist.org/observations/9869215
https://www.inaturalist.org/observations/3949369

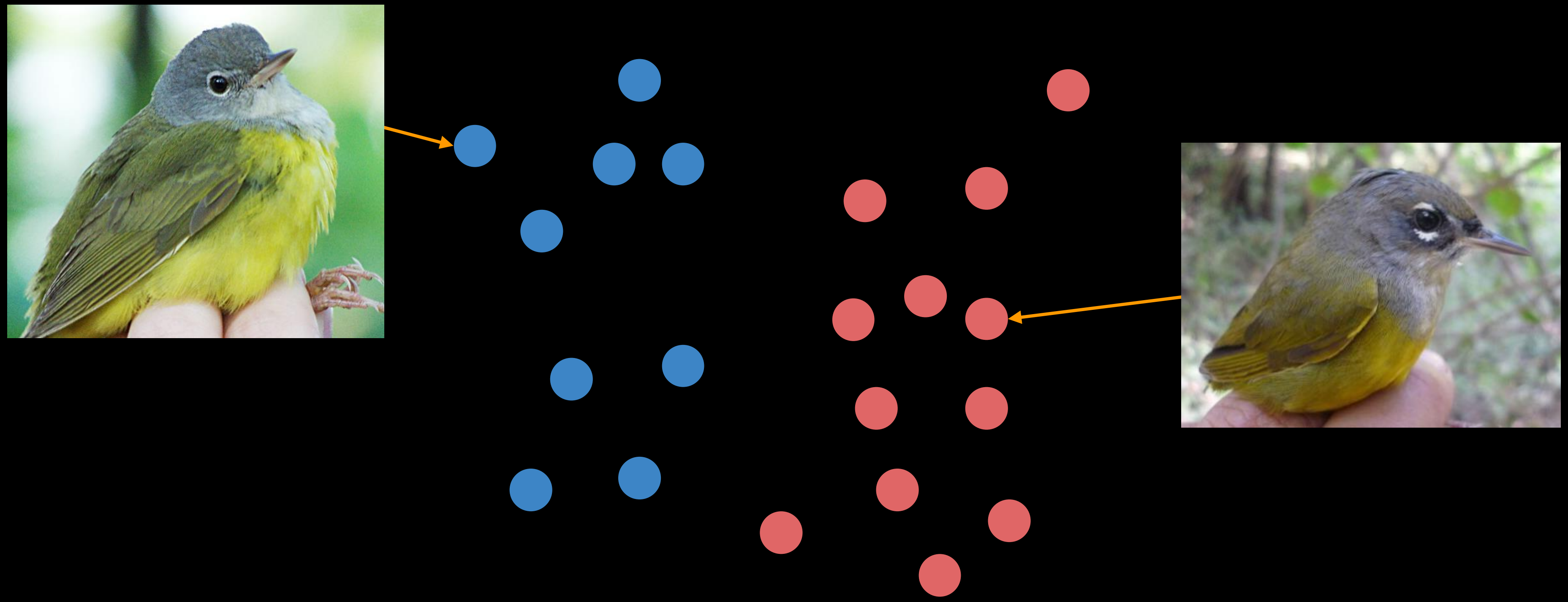
Connecticut Warbler

MacGillivray's Warbler

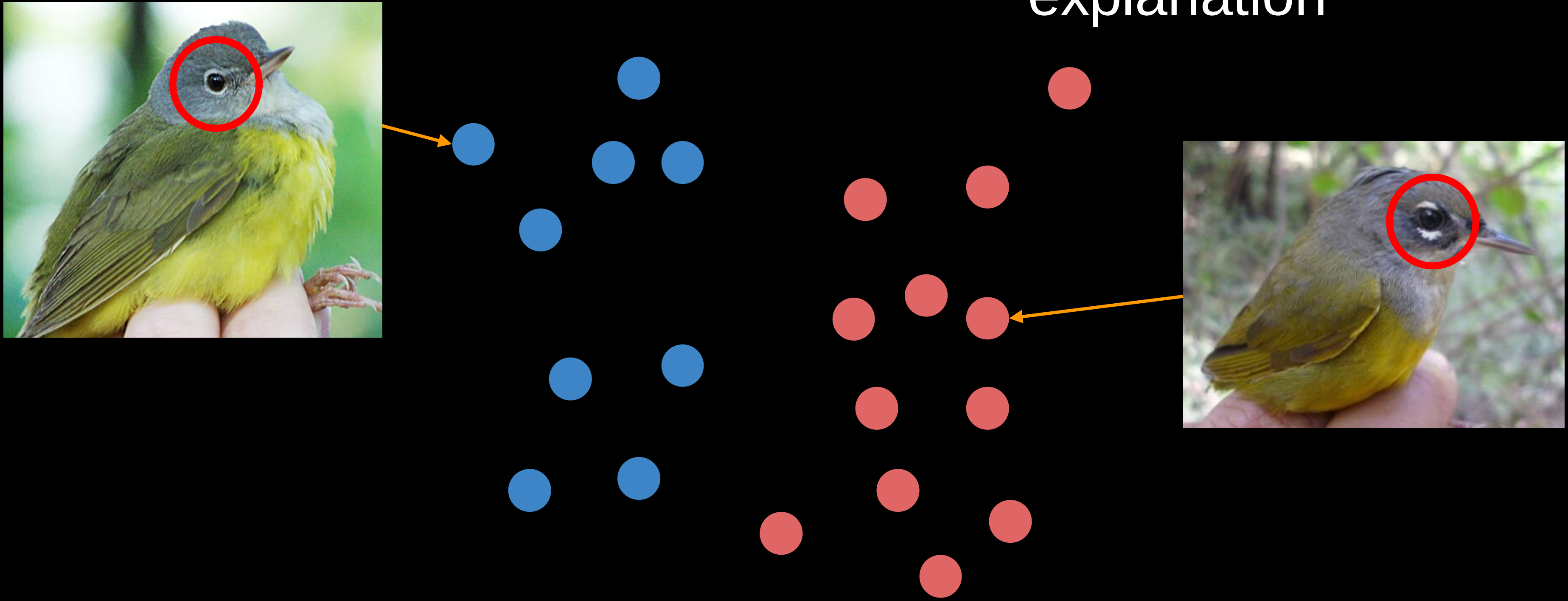
https://www.inaturalist.org/observations/9869215
https://www.inaturalist.org/observations/3949369

Teaching Categories to Human Learners with Visual Explanations **CVPR 2018**
Mac Aodha, Su, Chen, Perona, Yue, Singla

x is an image



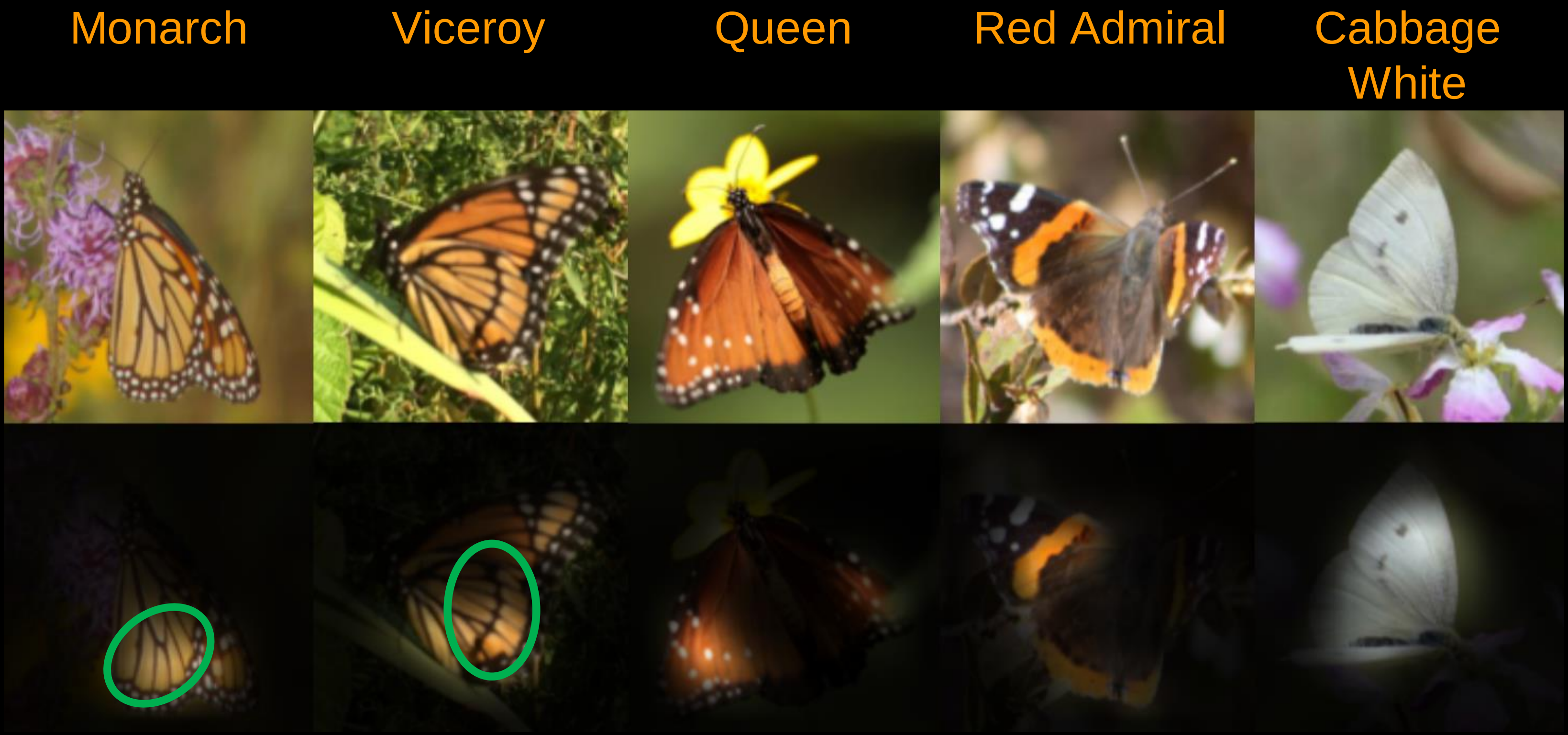
e is an associated explanation



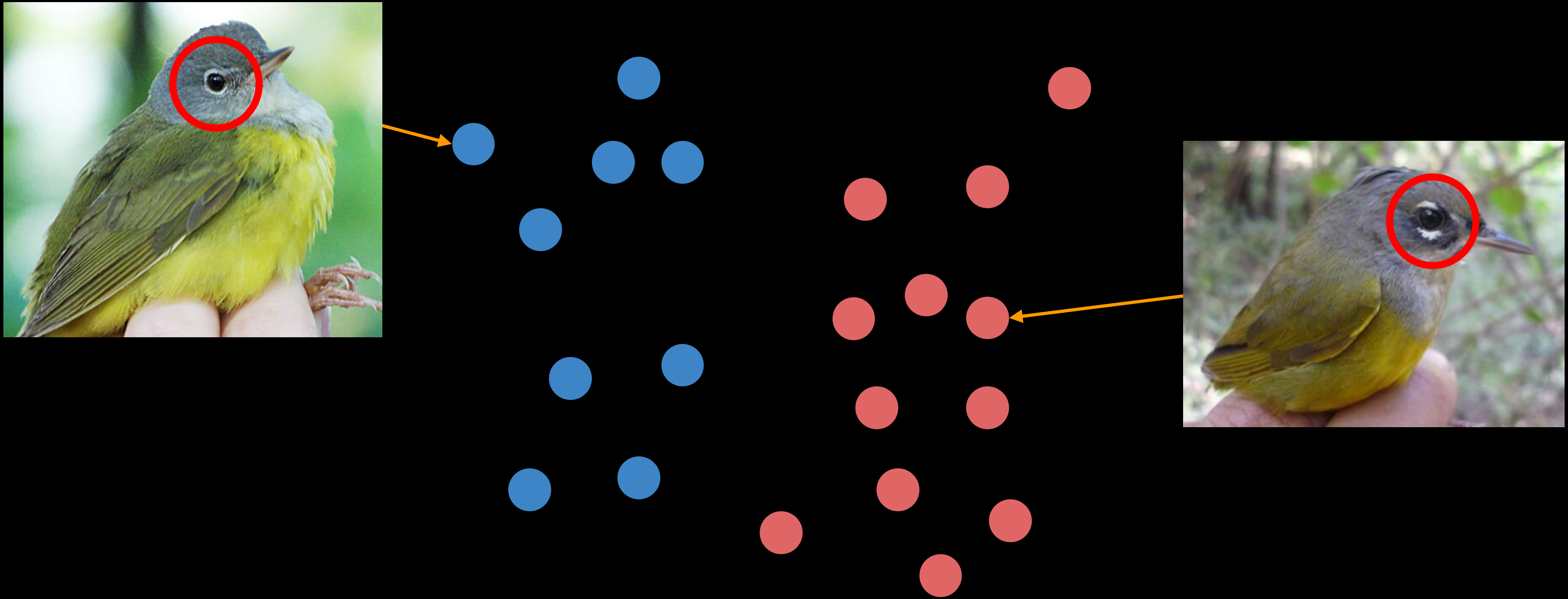
Butterflies

Monarch | Viceroy | Queen | Red Admiral | Cabbage White

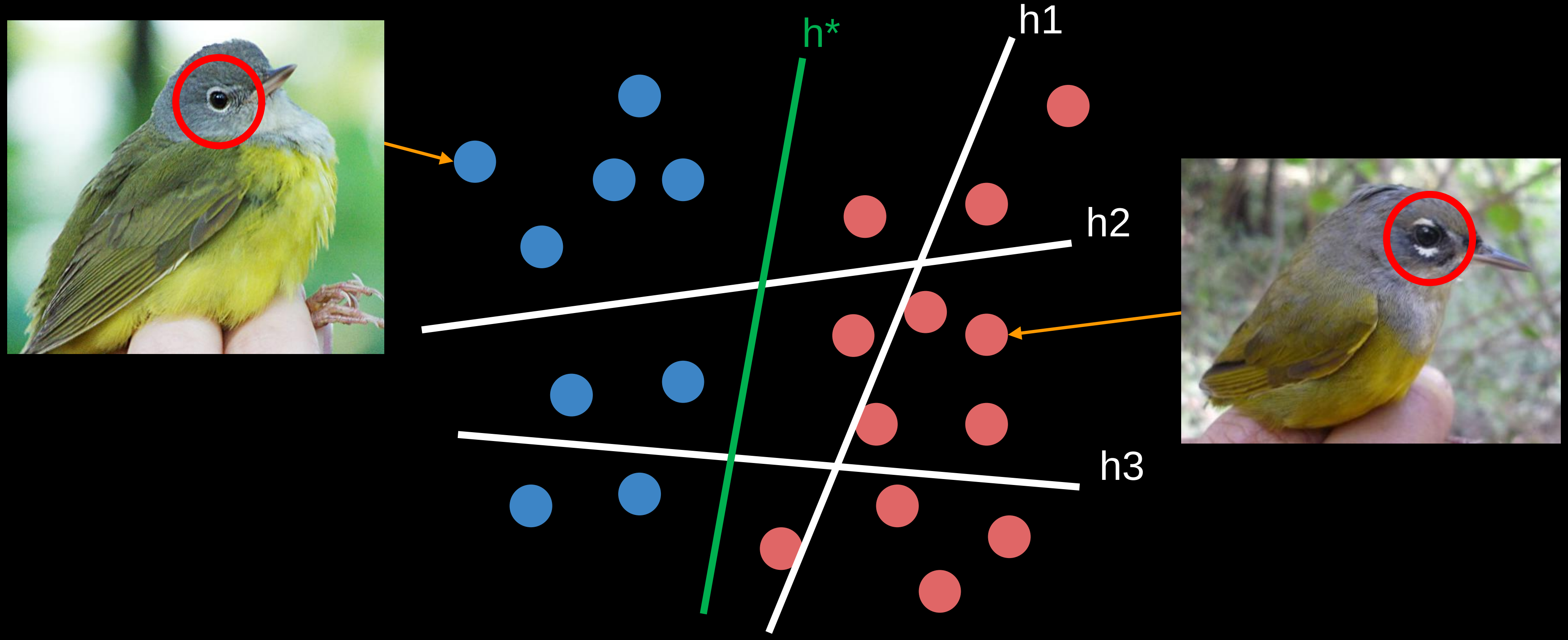
Butterflies

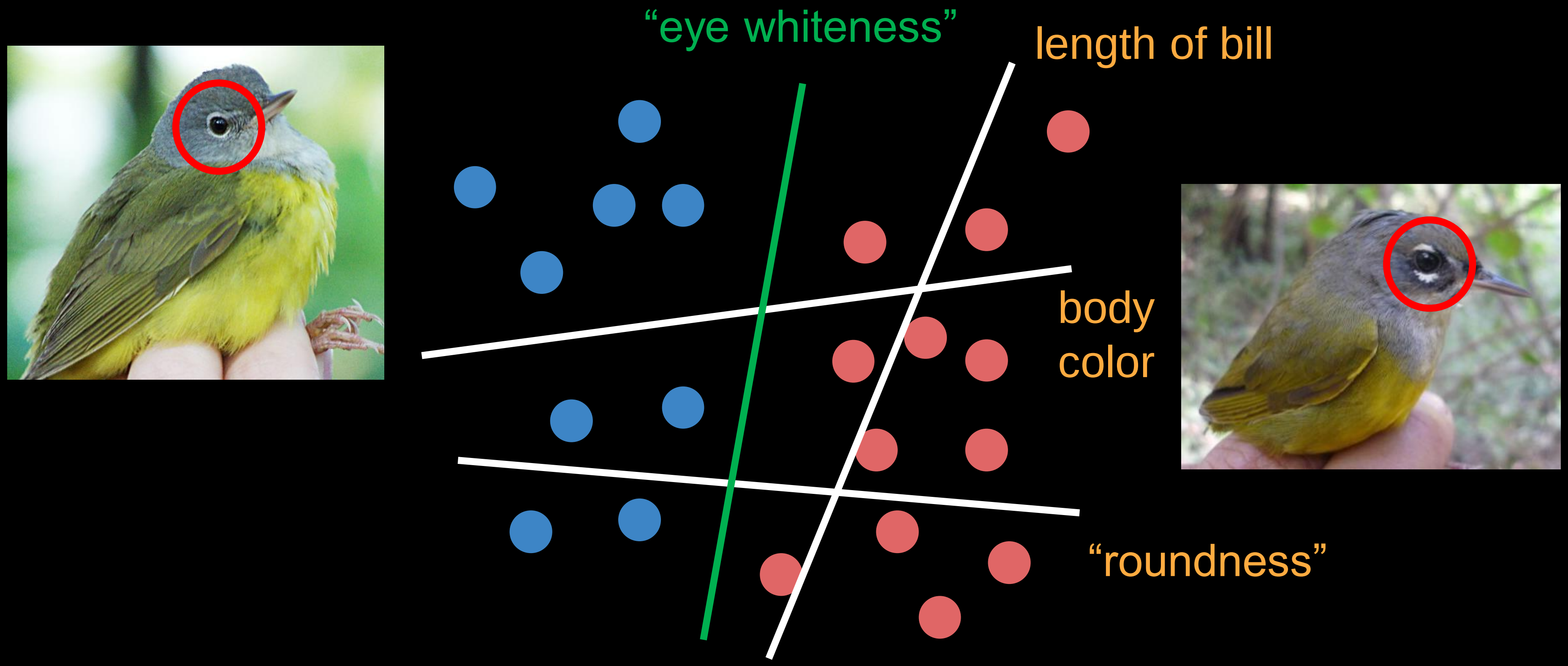


Learning Deep Features for Discriminative Localization **CVPR 2016**

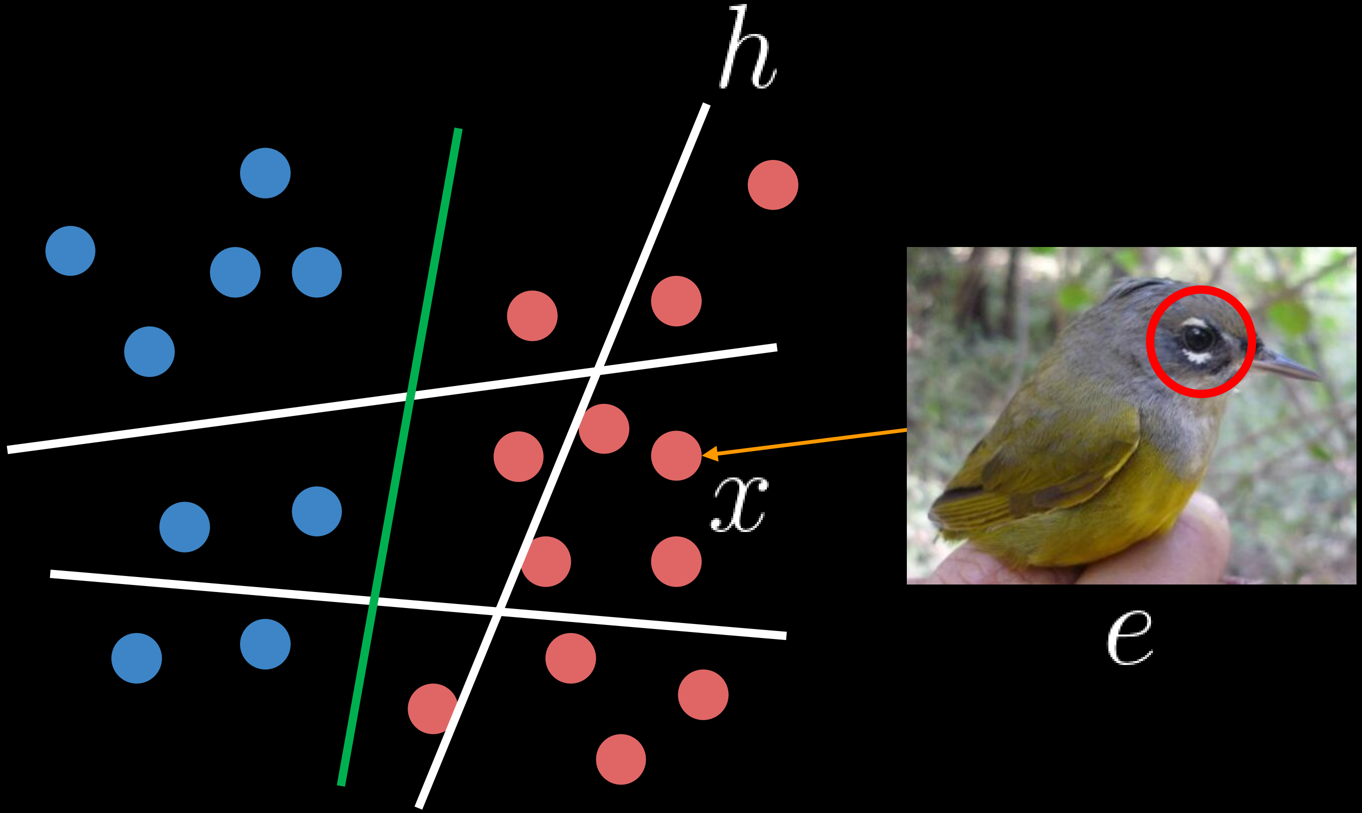


h is a hypothesis





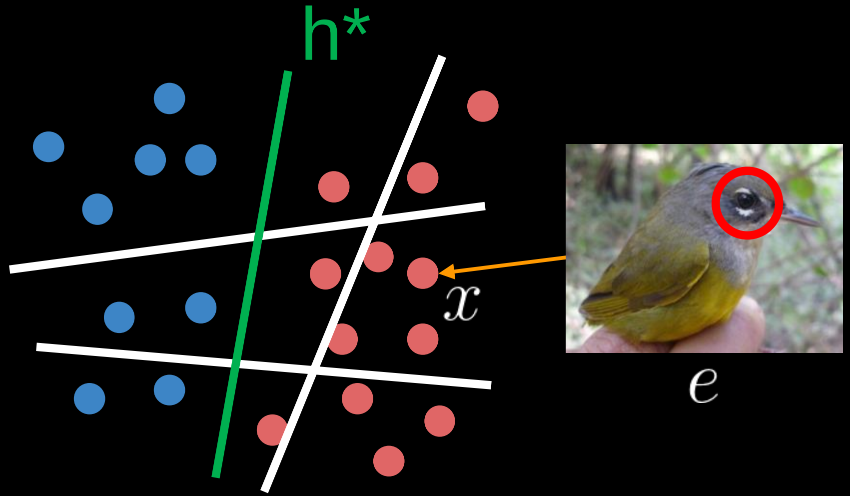
“eye whiteness”
length of bill
body color
“roundness”



h
x
e

How to Choose Teaching Set T to Teach h*?

$$T = \{(x_1, y_1, e_1), \ldots, (x_n, y_n, e_n)\}$$



Student Model

$$P(h|T)$$

Student Model

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t)$$

"win stay, lose switch"

Singla et al. Near-Optimally Teaching the Crowd to Classify **ICML 2014**

Student Model

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t)$$

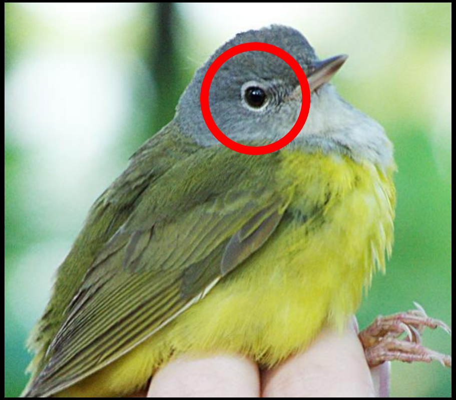
"win stay, lose switch"

$$S(y_t|h, x_t) = \begin{cases} 1 & \text{if} \quad y_t = \hat{y}_t^h \\ \frac{1}{1+\exp(-\alpha h(x_t) y_t)} & \text{otherwise} \end{cases}$$

Student Model - With Explanations

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t | h, x_t)$$

“Good”

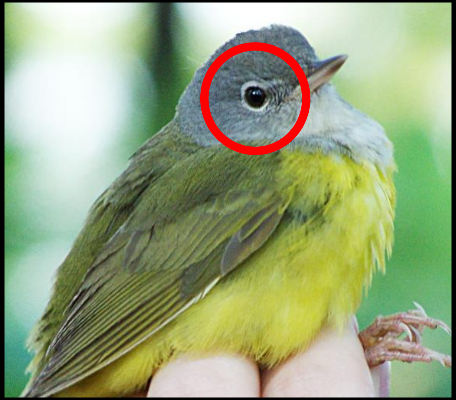


“Bad”

Student Model - With Explanations

$$P(h|T) \propto P(h) \prod_{x_t, y_t \in T} S(y_t|h, x_t) \prod_{x_t, e_t \in T} \boxed{(E(e_t)} D(x_t))$$

"Good"



"Bad"

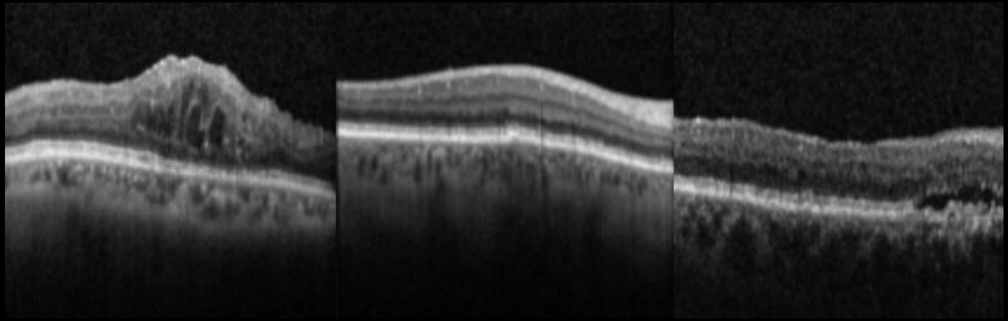
Selecting the Teaching Set *T*

Select for largest reduction in expected error

$$\mathbb{E}[err(h)|T] = \sum_{h \in \mathcal{H}} P(h|T)err(h)$$

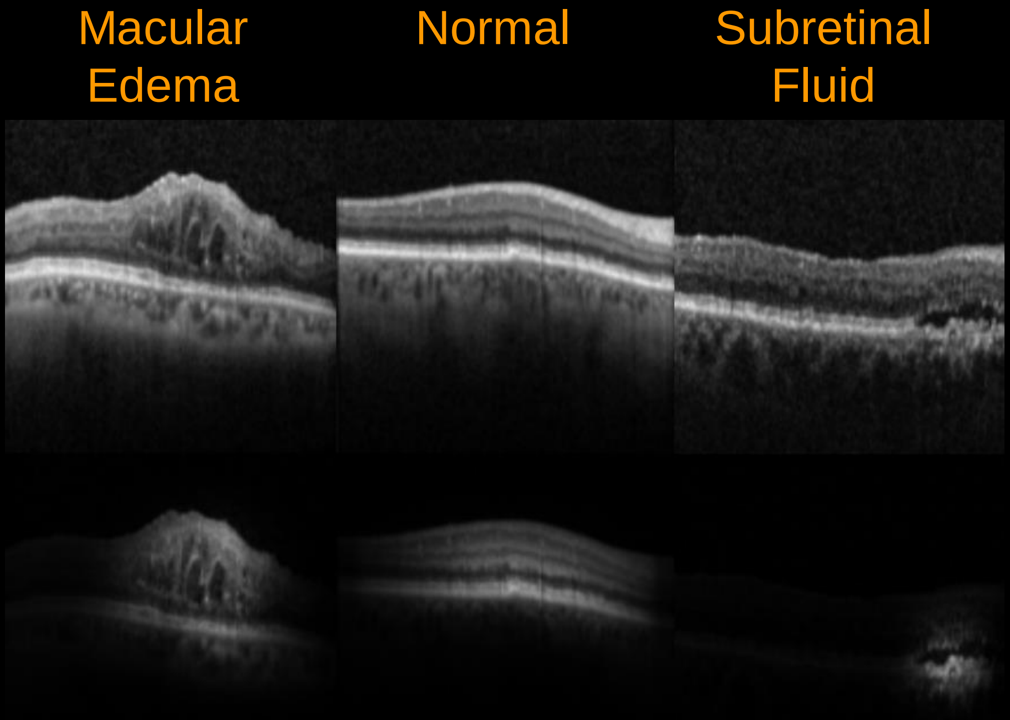
Retina Images

1125 images, 3 classes



Retina Images

1125 images, 3 classes



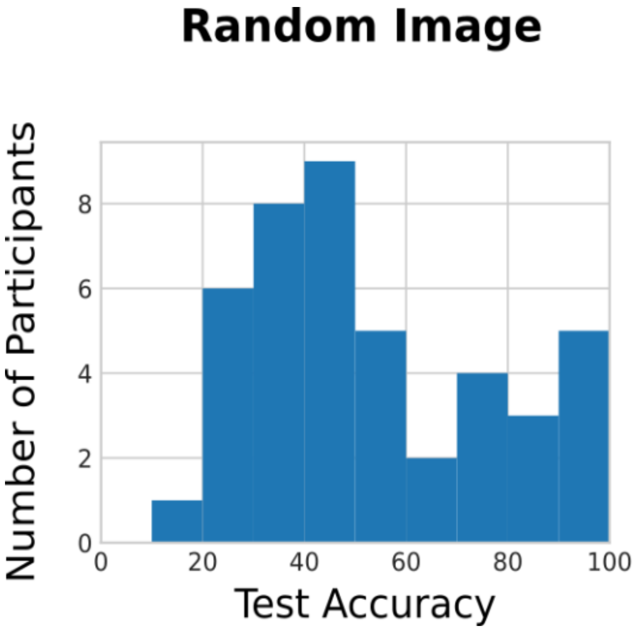
Experimental Setup

Familiarize participants with interface

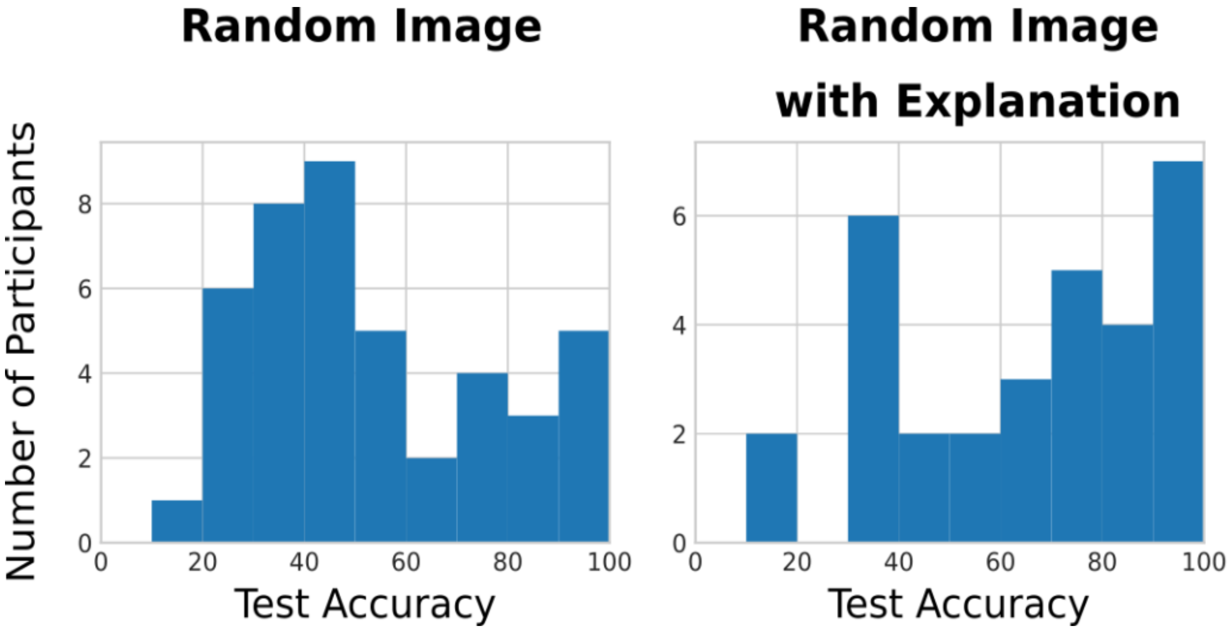
Teach for 20 iterations

Test for 20 iterations (to measure performance)

Results for Retina Images



Results for Retina Images



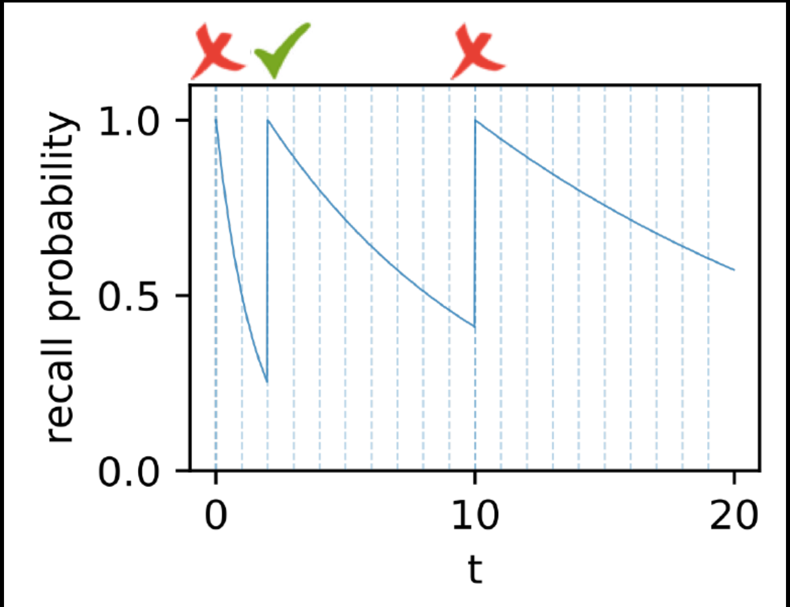
Results for Retina Images

Modeling Learner Memory Decay

Memory decays over time

Spaced repetition model

$P_i(t \mid \text{history})$



Teaching Multiple Concepts to Forgetful Learners **NeurIPS 2019**

Hunziker, Chen, **Mac Aodha**, Gomez Rodriguez, Krause, Perona, Yue, Singla

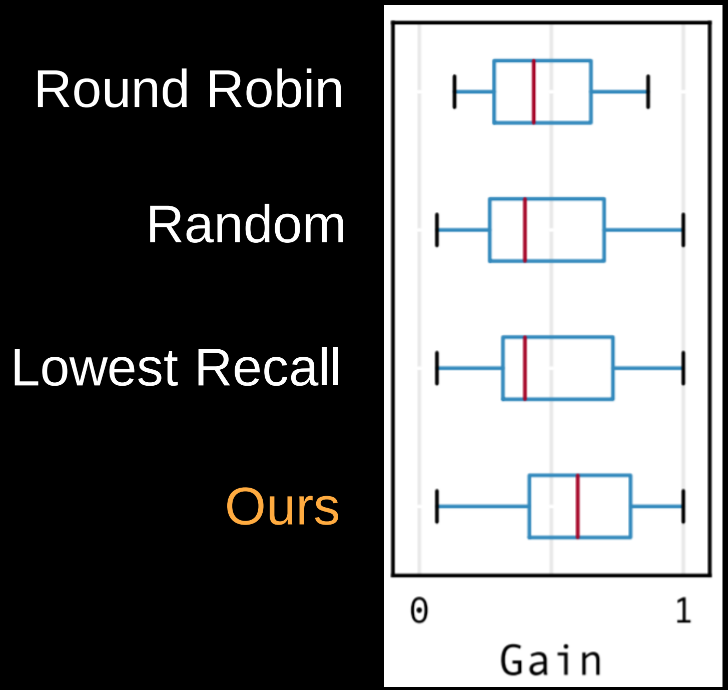
Modeling Learner Memory Decay

“Flash cards”

15 concepts, 40 time steps

~ 25 mins per exp

German vocabulary



Teaching Multiple Concepts to Forgetful Learners **NeurIPS 2019**
Hunziker, Chen, **Mac Aodha**, Gomez Rodriguez, Krause, Perona, Yue, Singla

Open Questions

- Models of the learning process
- Learning to teach
- Learning how low shot is performed
- Evaluation tool for interpretable machine learning
- Generating curriculums

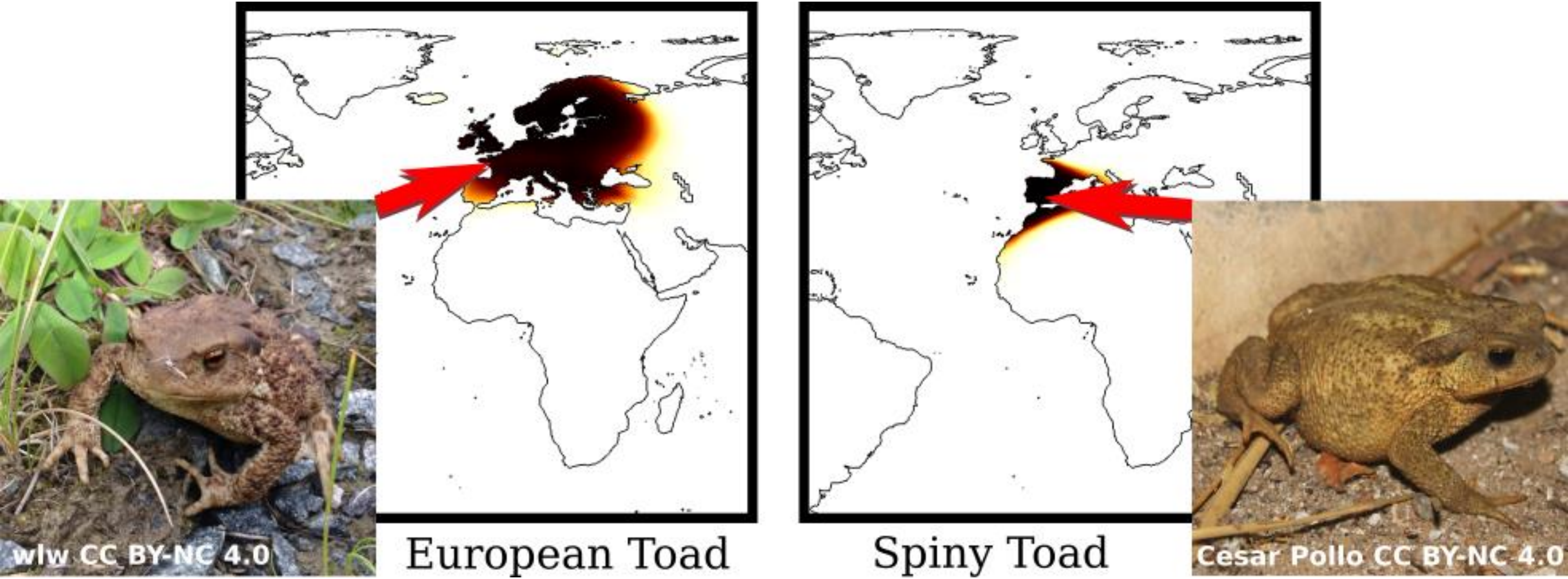
Presence-Only Geographical Priors for Fine-Grained Image Classification **ICCV 2019**
Mac Aodha, Cole, Perona

Can we use information such as where, when, and who captured an image to help determine its class?

Presence-Only Geographical Priors for Fine-Grained Image Classification **ICCV 2019**
Mac Aodha, Cole, Perona

Which class *y* is in image *I* ?

$$P(y|I)$$


wlw CC BY-NC 4.0
European Toad
Spiny Toad
Cesar Pollo CC BY-NC 4.0

Which class y is in image I taken at location $\mathbf{x}$?

$$P(y|I, \mathbf{x})$$

Which class y is in image I taken at location $\mathbf{x}$?

$$P(y|I,\mathbf{x}) \propto P(y|I)P(y|\mathbf{x})$$

Which class y is in image I taken at location $\mathbf{x}$?

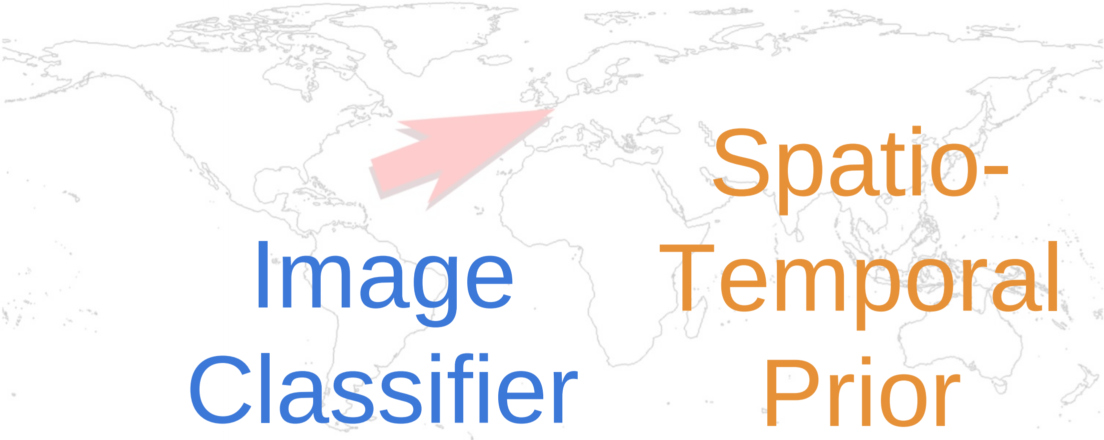


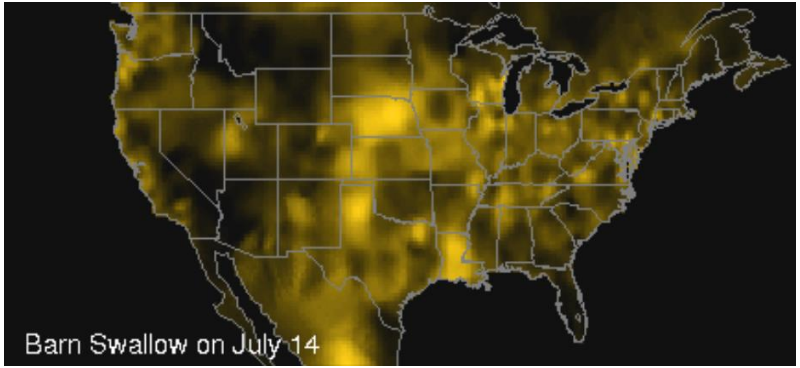
Image Classifier

Spatio-Temporal Prior

$$P(y|I,\mathbf{x}) \propto \boxed{P(y|I)}\boxed{P(y|\mathbf{x})}$$

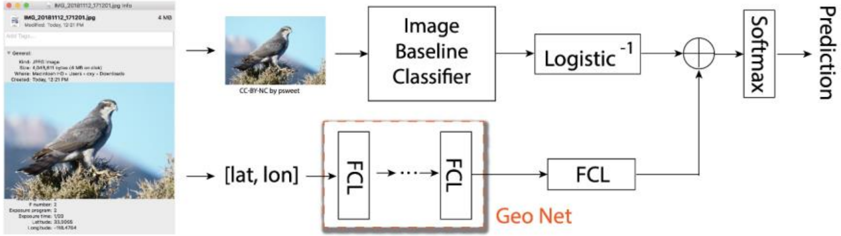
Previous Work

“Non-parametric”

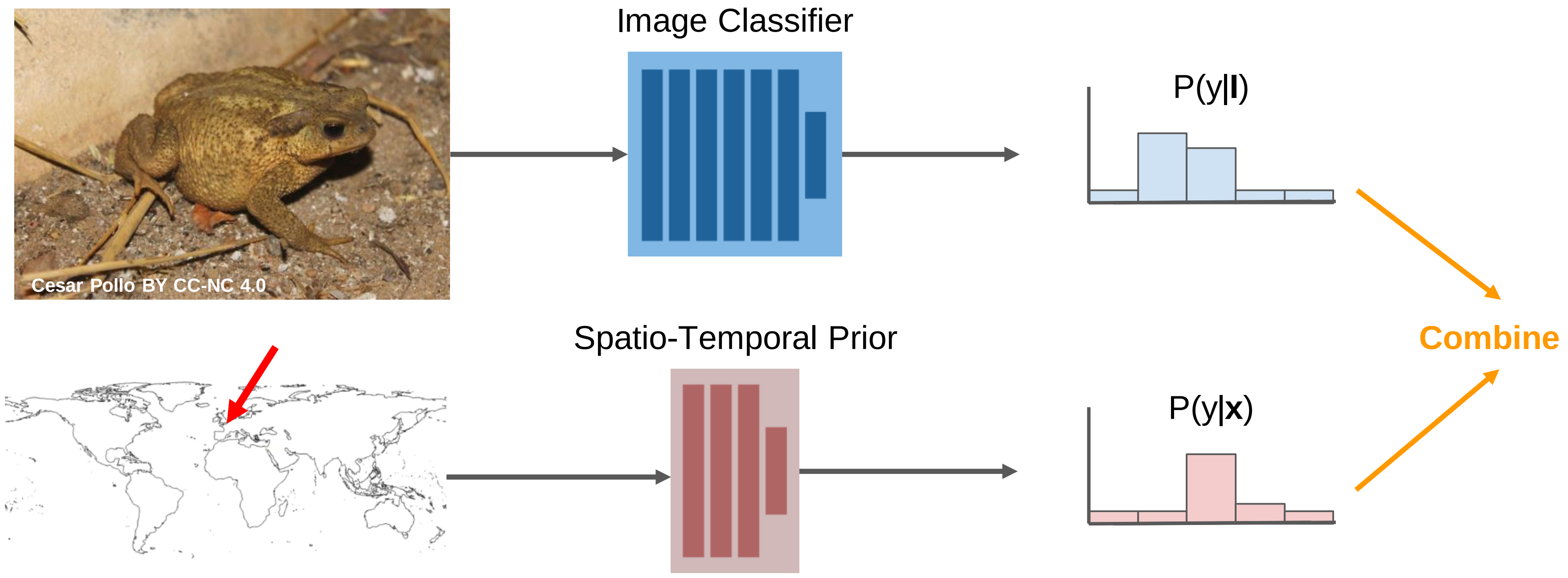


Berg et al. CVPR 2014

Jointly Trained



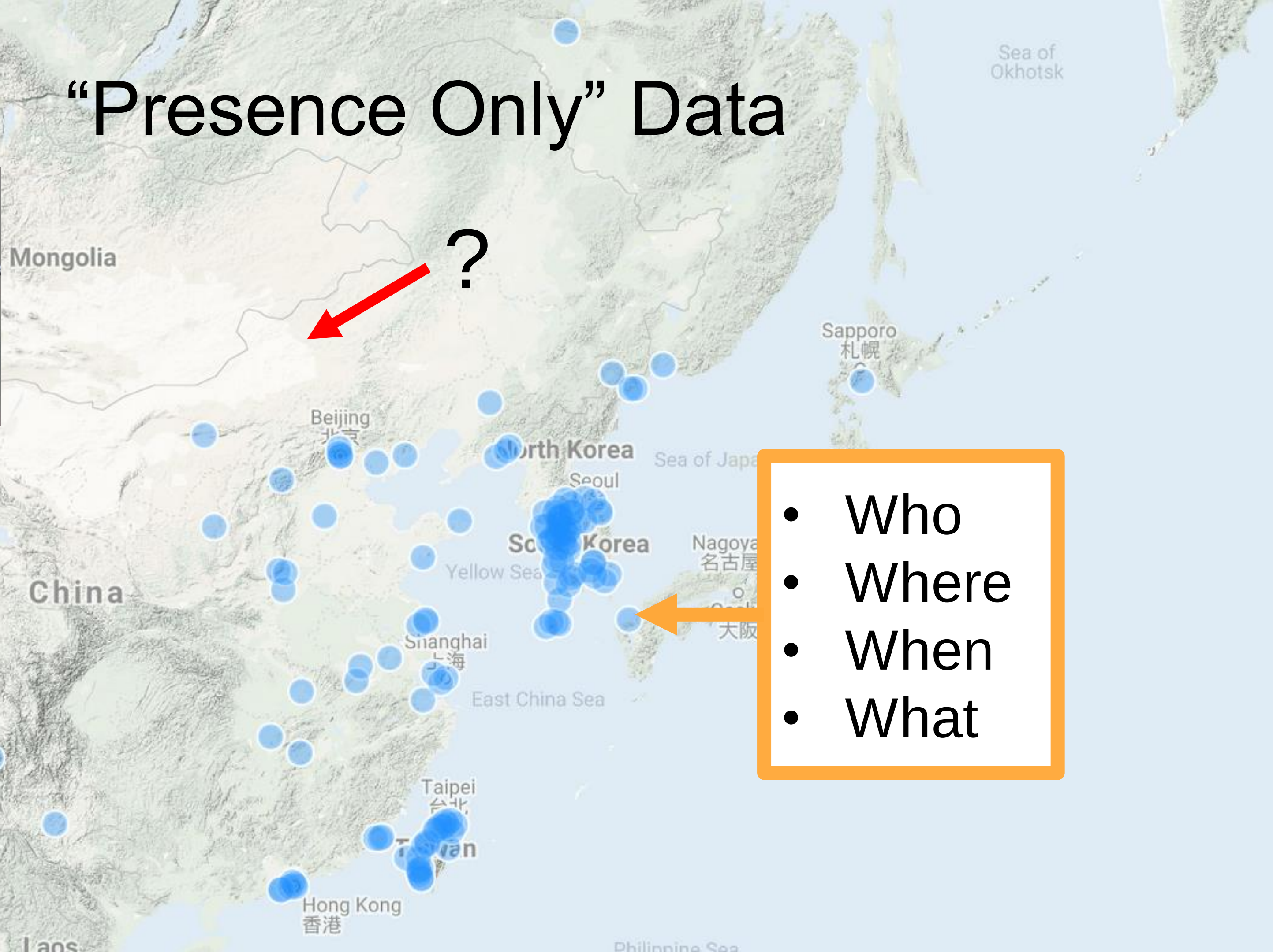
Tang et al. ICCV 2015
Chu et al. Arxiv 1906.01737, 2019



x = (longitude, latitude, day)

“Presence Only” Data

Oriental Magpie



- Who
- Where
- When
- What

Training Data

location (+time) **photographer ID**

$$\mathcal{D} = \{(\mathbf{x}_i, y_i, p_i) | i = 1, ..., N\}$$

class label

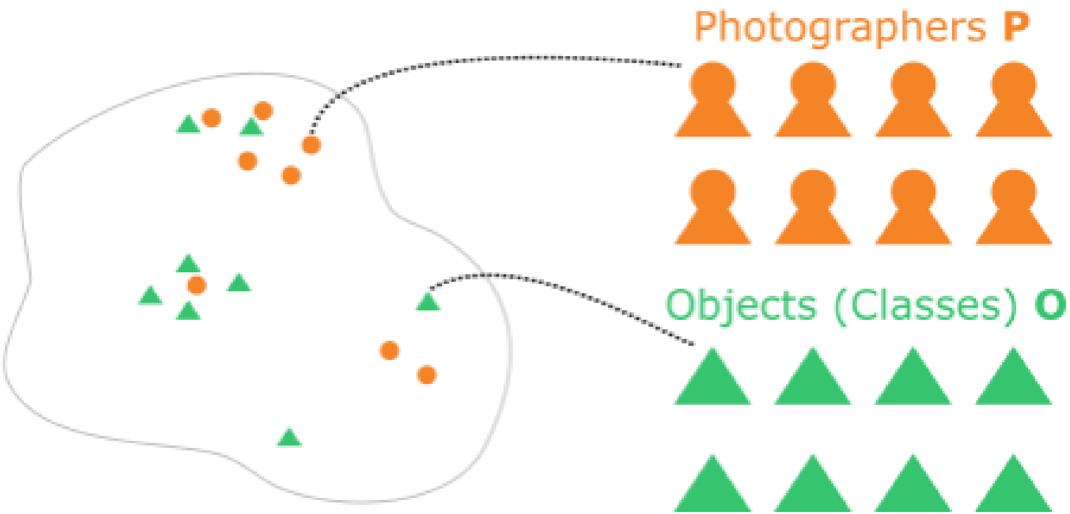
Joint Model of Photographers and Objects

Capture the following relationships:

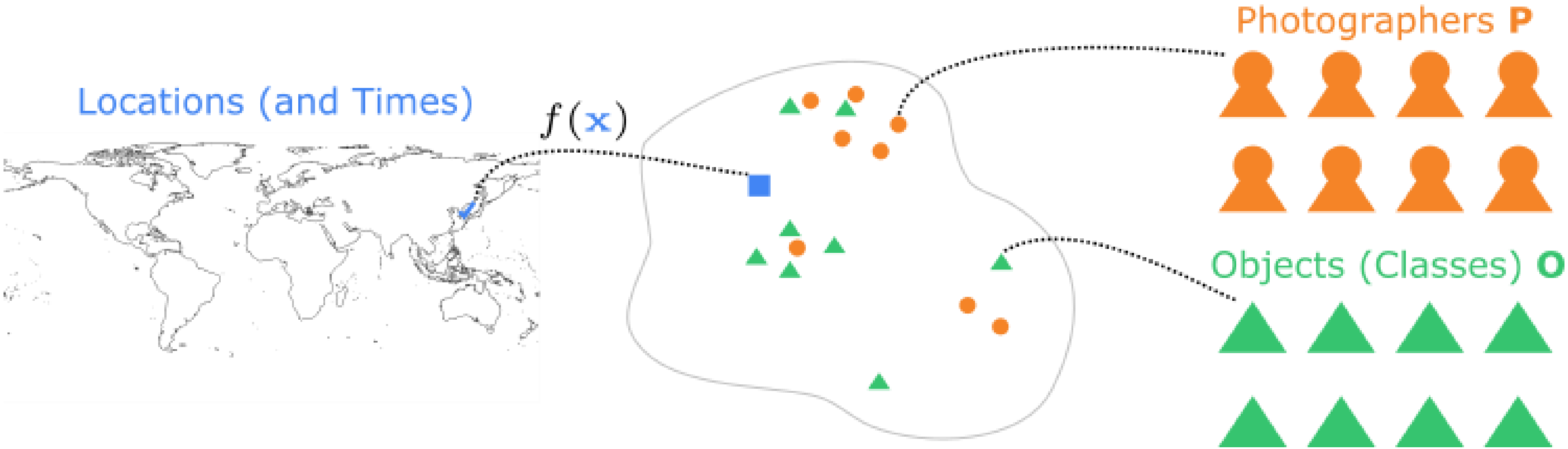
- class X’s affinity for location Y
- photographer Z’s affinity for class X
- photographer Z’s affinity for location Y

Shared Embedding Space

Shared Embedding Space

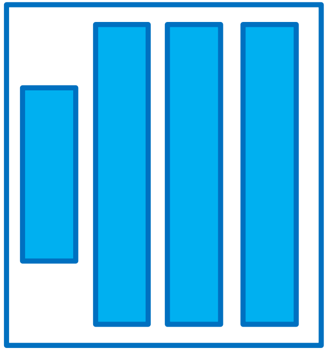


Shared Embedding Space

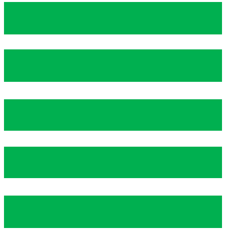


Learned Parameters

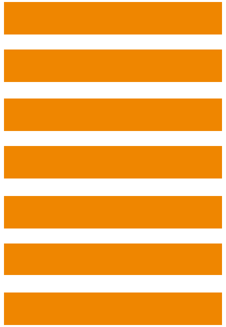
$$\theta = [\theta_f, \mathbf{O}, \mathbf{P}]$$



Location
Embedder **f()**



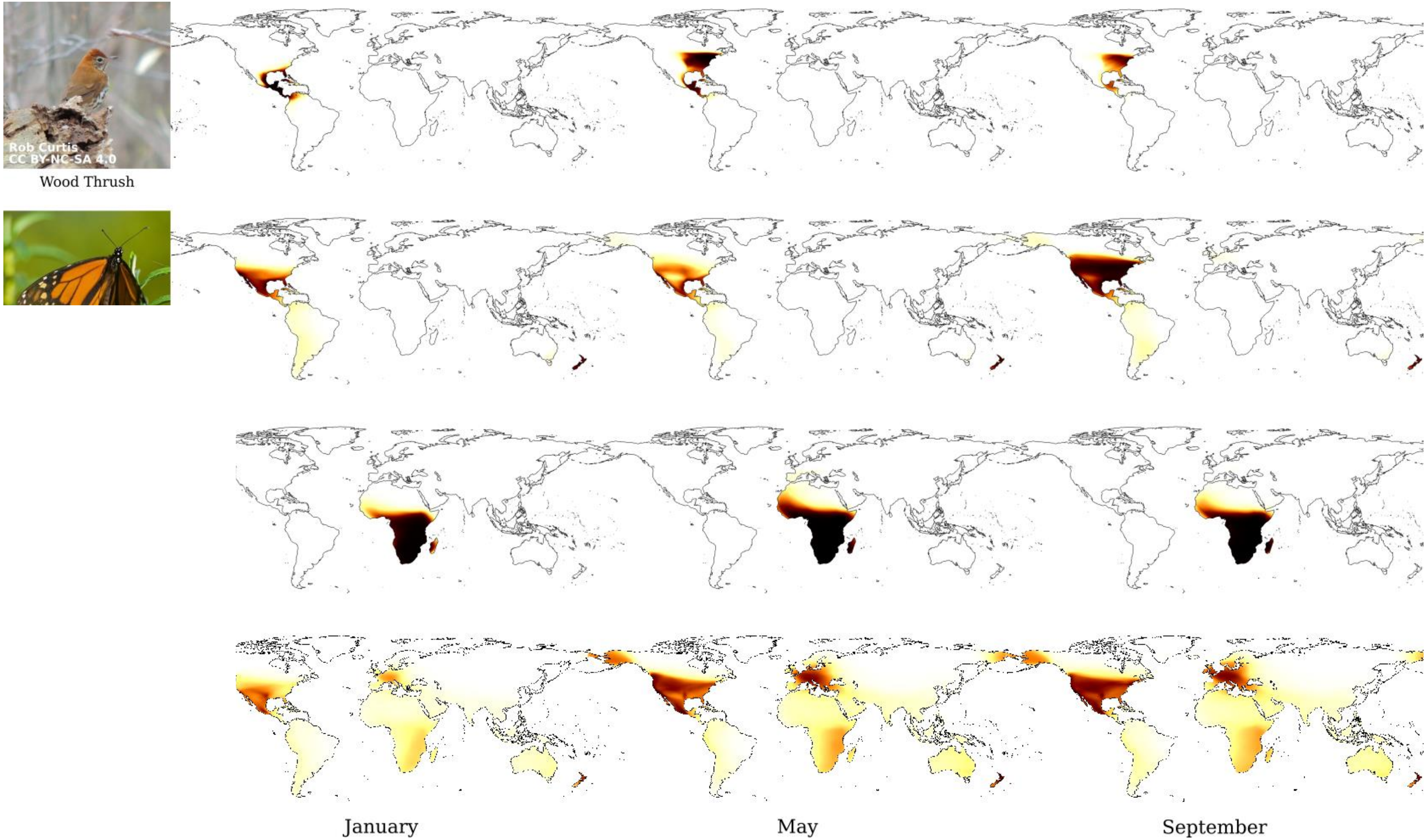
Class
Embeddings **O**



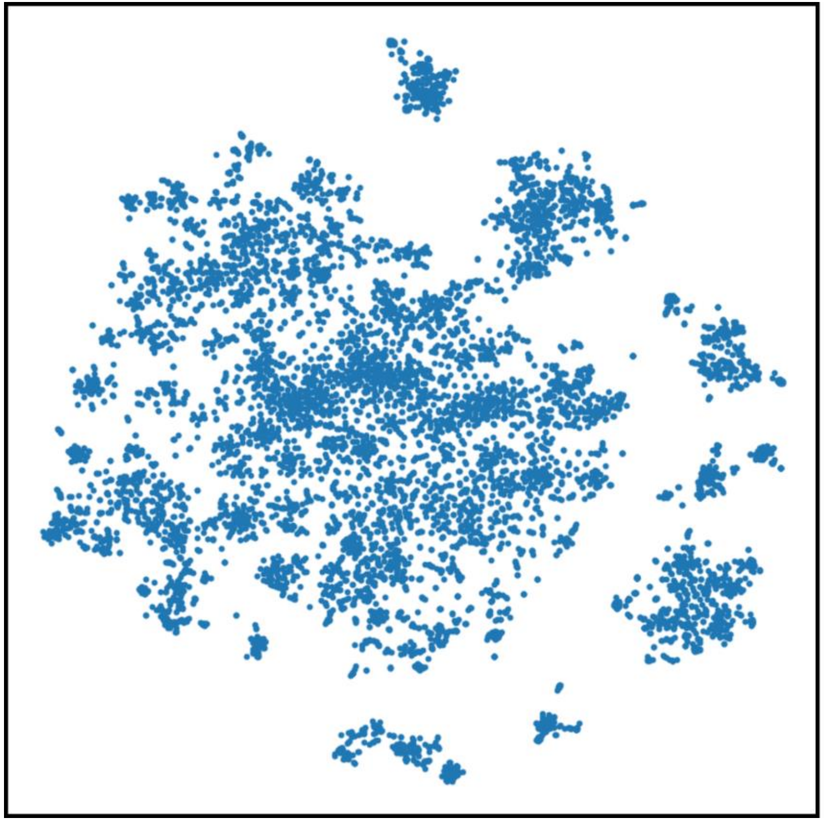
Photographer
Embeddings **P**

Wood Thrush

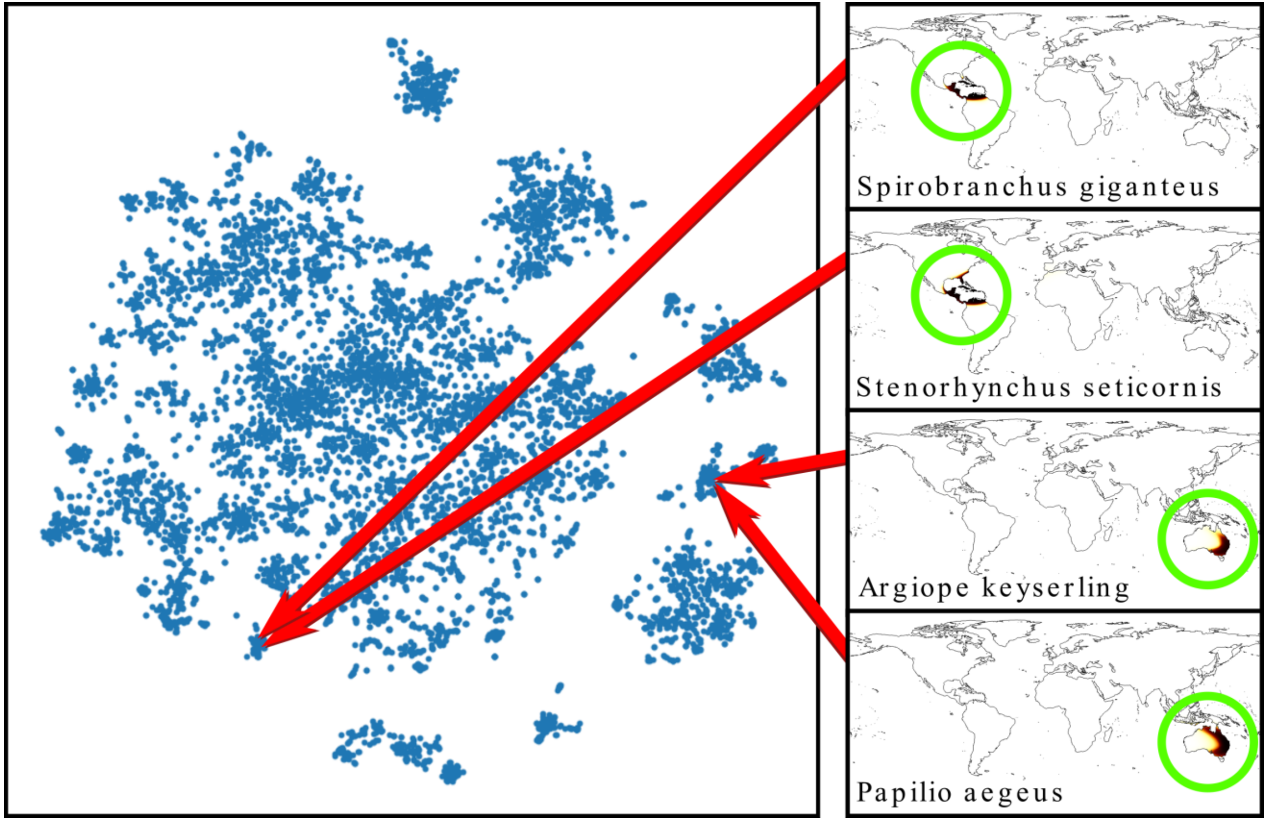
Predicted Locations


Rob Curtis
CC BY-NC-SA 4.0
Wood Thrush
January
May
September

Category Embedding **O**



Category Embedding **O**

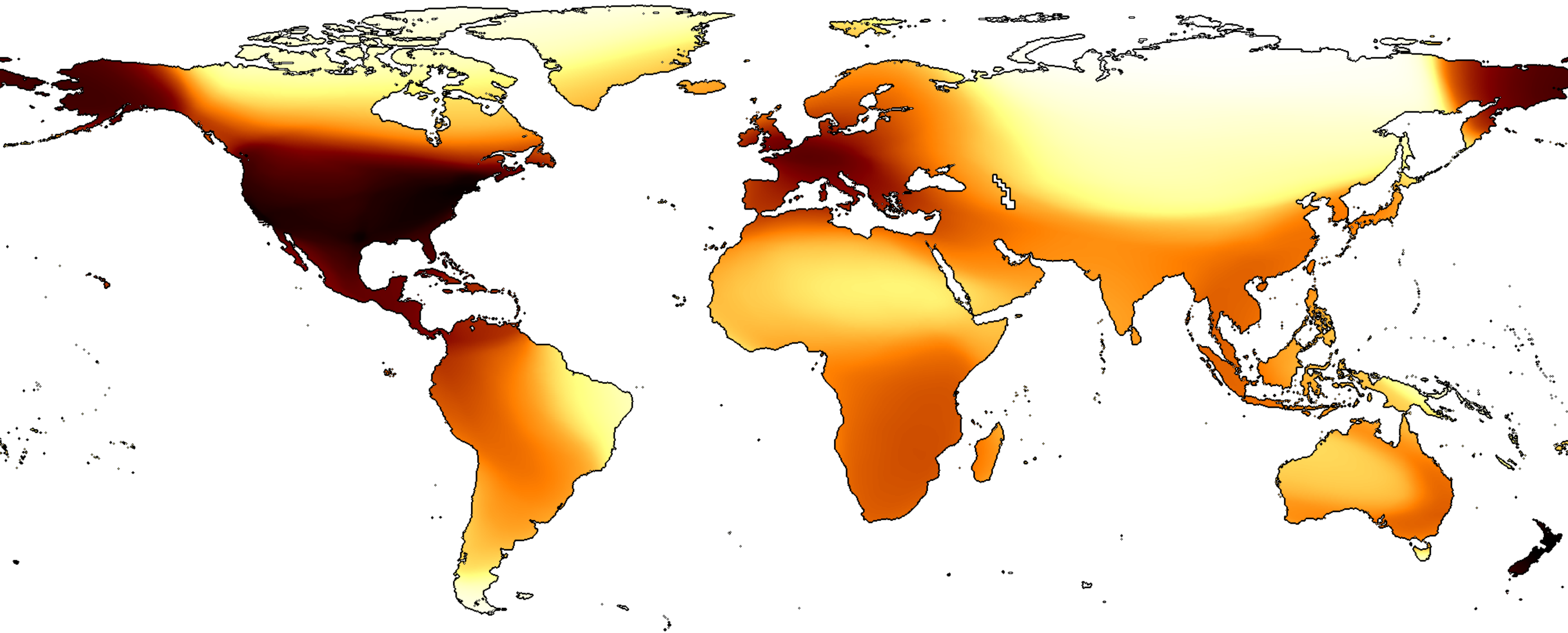


Visualization of Spatio-Temporal Predictions

Embedding of Each Location on the Earth

Location
Embedder **f()**

Photographer Location Affinity



Photographer Embedding **P**

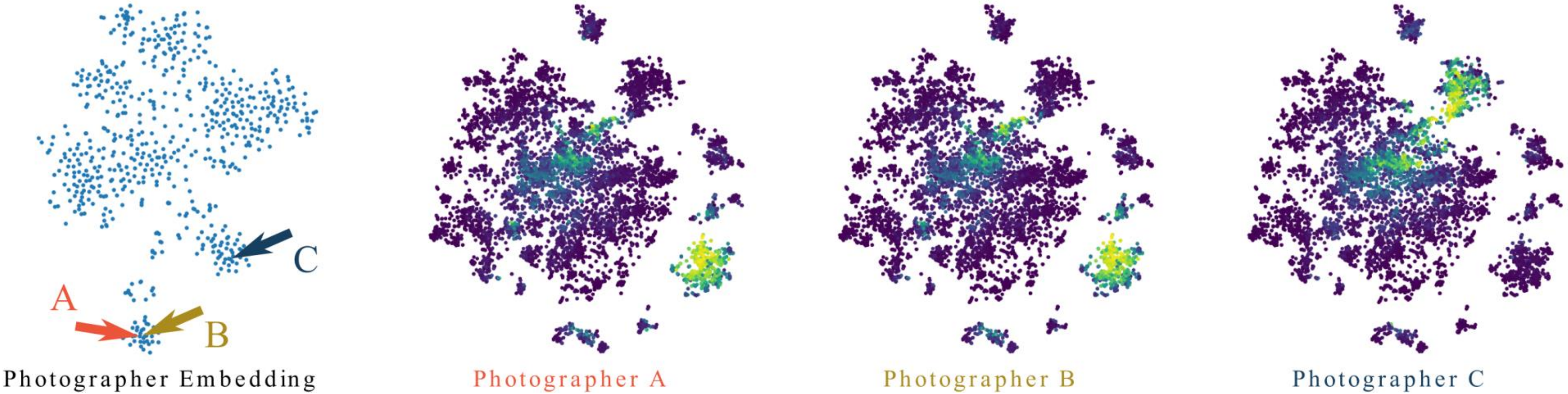
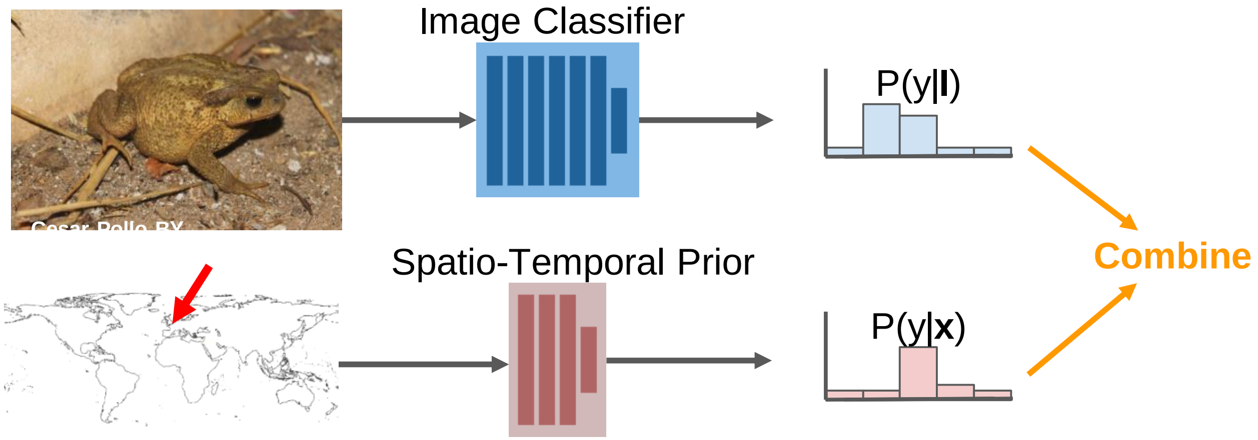
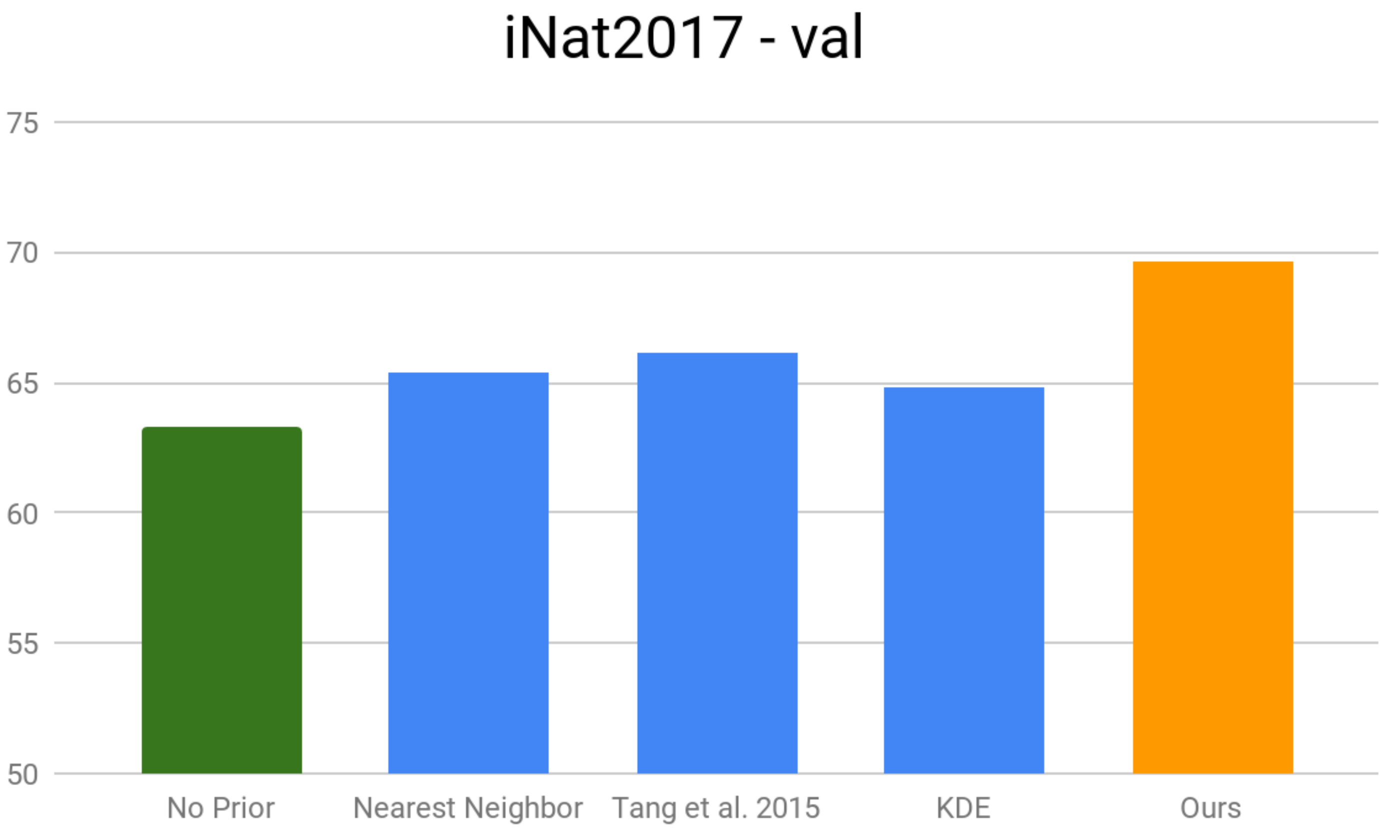


Image Classification Results

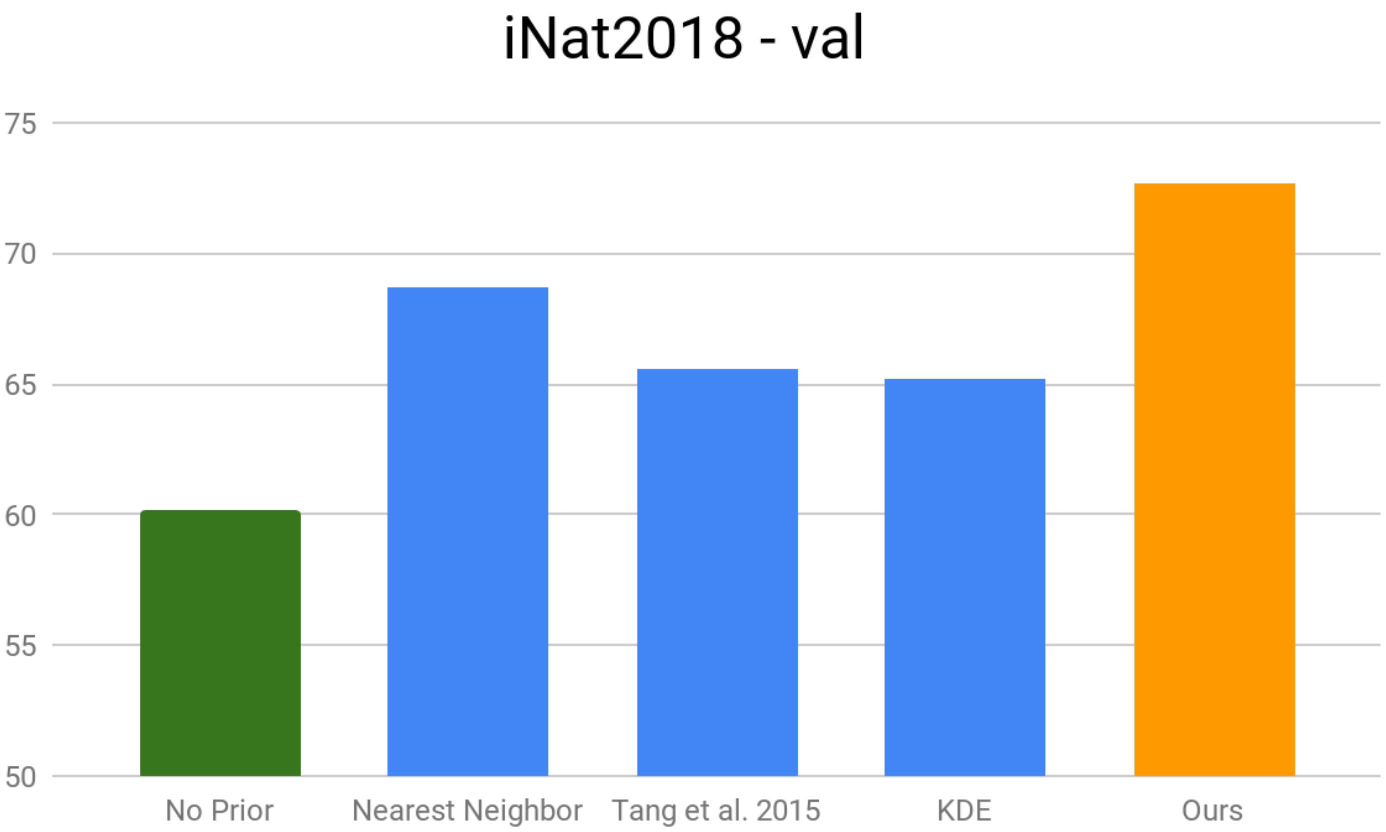
Can our prior improve image classification performance?



x = (longitude, latitude, day)



$$P(y|I) \qquad\qquad P(y|I, \mathbf{x})$$



$$P(y|I) \qquad\qquad P(y|I, \mathbf{x})$$

Hylocichla mustelina - Wood Thrush

Jan

Type the name of a particular species or click "random".

Search..

search random

About

<u>Today 3-4:30pm</u>

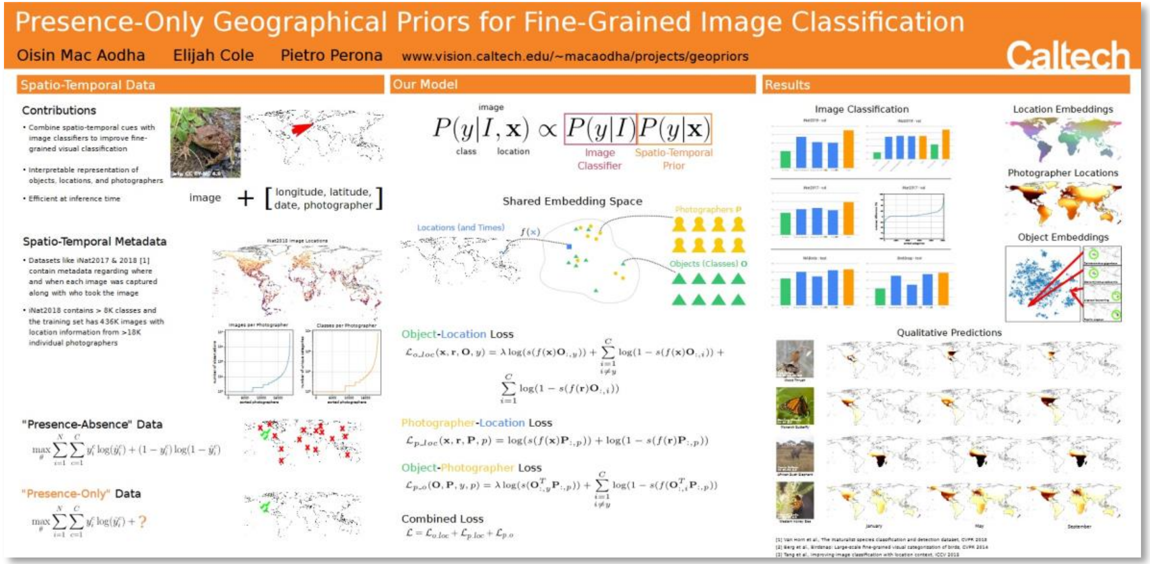
Poster #140

<u>Friday, Nov 1st</u>

Poster #45

15:30–18:00

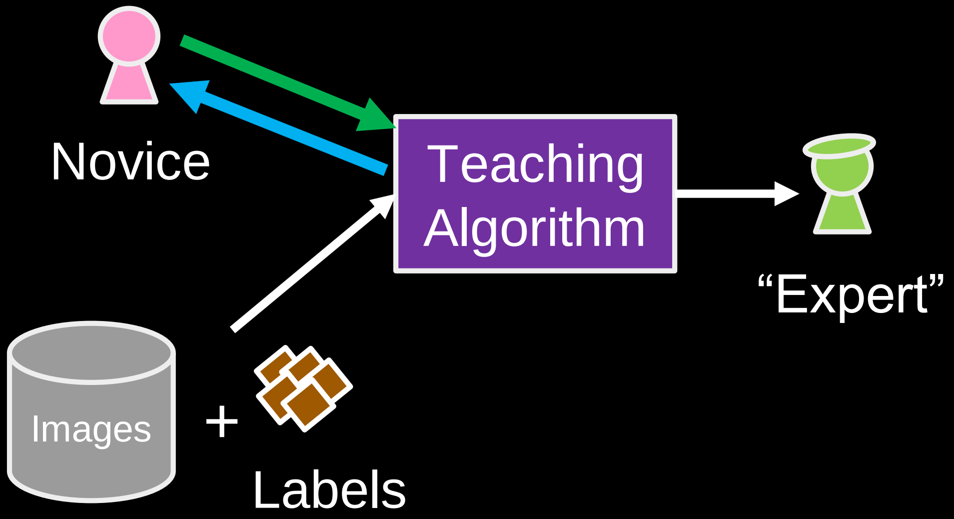
Poster 4.2 (Hall B)



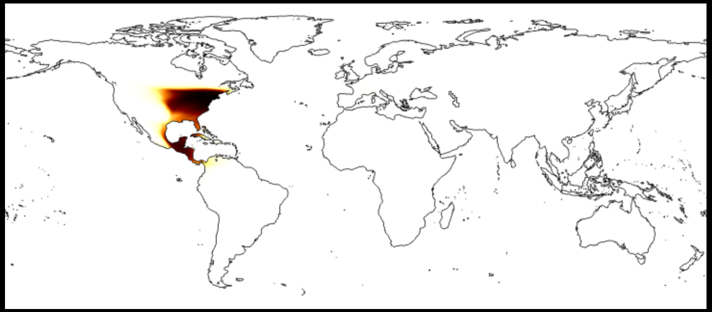
www.vision.caltech.edu/~macaodha/projects/geopriors

Improving Fine-Grained Classifiers

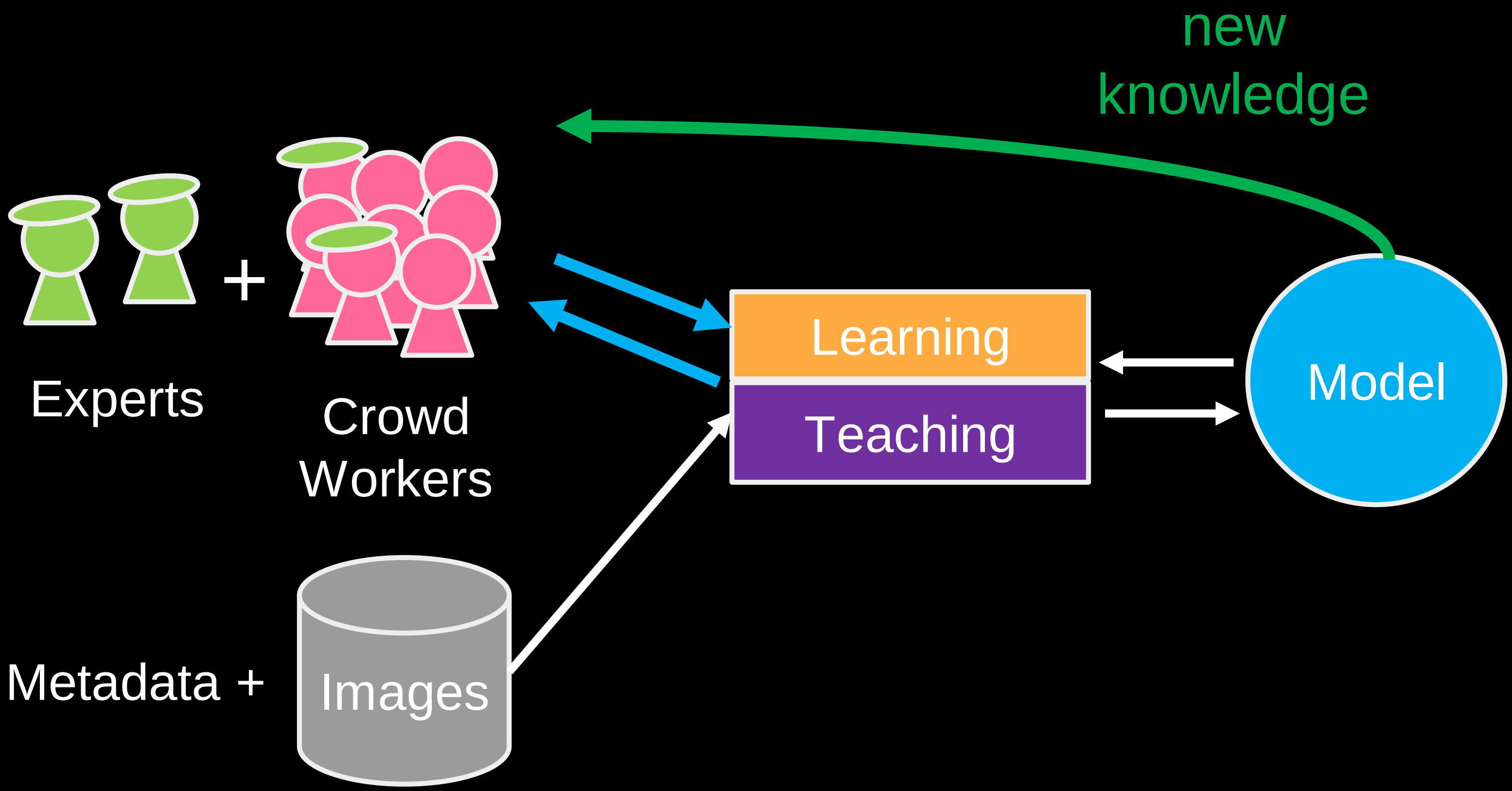
More Annotations



Metadata



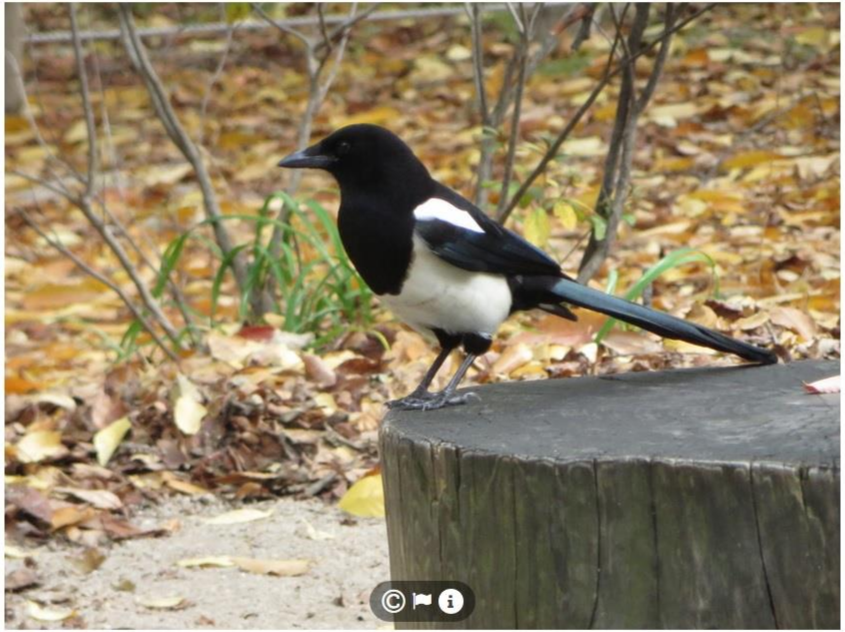
Humans-in-the-Loop



iNaturalist Search Explore Your Observations Community Identify More Upload 0 0

Oriental Magpie (Pica serica) Research Grade

Follow



Male/Female

Life Stage

Behavior

Health

Individual ID

…

Activity

dmkidwell suggested an ID Improving 3d

Oriental Magpie

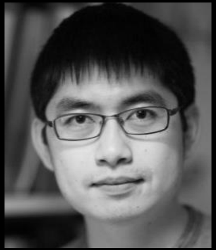
Community Taxon What's this?

Oriental Magpie (Pica serica)

Cumulative IDs: 4 of 4

Thanks!

Elijah Cole
Caltech



Yuxin Chen
Uni. of Chicago

Grant Van Horn
Cornell

Yisong Yue
Caltech

Serge Belongie
Cornell

Pietro Perona
Caltech

More info at www.oisin.info